

=> d his

(FILE 'REGISTRY' ENTERED AT 08:37:37 ON 15 JUL 2001)

DEL HIS

E PCAPGTFSNTTSSTDICRPHQIC/SQEP

E PCAPGTFSNTTSSTDICRPHQICNVVALIPGNASMDAVCT/SQEP

E TSTSPTRSMAPGAVHLP/SQEP

E STSPTRSMAPGAVHLP/SQEP

No hits on closed sequences

GenCore version 4.5 -
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:15:46 ; Search time 12.7 Seconds
(without alignments)
137.954 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_185
Perfect score: 134
Sequence: 1 PCAPGTFSTNTSTDICRPHQIC 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4466

Minimum DB seq length: 0
Maximum DB seq length: 23

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	23.1	20	2 A42865	Ca2+/calmodulin-de
2	30	22.4	18	2 B49048	T-cell receptor be
3	29	21.6	13	2 S47381	T-cell antigen rec
4	29	21.6	16	2 S38292	30K allergen - rye
5	29	21.6	17	2 A60317	glucagon-like pept
6	29	21.6	21	2 I34351	gene HEXA protein
7	29	21.6	22	2 A39269	Lx-1 tumor antigen
8	28	20.9	23	2 S60565	homeodomain protei
9	27	20.1	20	2 T48881	leader peptide [im
10	27	20.1	22	1 MXKN1	mu-conotoxin GIIIA
11	26	19.4	15	2 JN0263	antigen (clone PVL
12	26	19.4	16	2 B60278	24K antigen - Myco
13	26	19.4	20	2 I54283	arylsulfatase A -
14	26	19.4	23	2 A59048	convulsant peptide
15	25.5	19.0	22	2 I37144	aspartylglycosamin
16	25.5	19.0	23	2 E39855	paralytic peptide
17	25.5	19.0	23	2 D39855	paralytic peptide
18	25	18.7	11	2 S23308	substance P - rain
19	25	18.7	13	2 PQ0491	self-incompatibili
20	25	18.7	13	2 D56661	S-locus specific g
21	25	18.7	17	2 B61334	trypsin (EC 3.4.21
22	25	18.7	19	2 B56613	virion morphogenes
23	25	18.7	19	2 C56661	S-locus specific g
24	25	18.7	19	2 PQ0492	self-incompatibili
25	25	18.7	22	1 MXKN2	mu-conotoxin GIIIB
26	25	18.7	22	1 MXKN3	mu-conotoxin GIIIC
27	25	18.7	23	2 A48968	exo-poly-alpha-gal
28	25	18.7	23	2 B38671	peptidylglycine mo
29	24	17.9	11	2 PH0891	T-cell receptor be

30	24	17.9	14	2 PH0776	T-cell receptor al
31	24	17.9	19	2 A05305	hemoglobin beta-2
32	24	17.9	21	2 C39543	collagen alpha 3(I
33	23.5	17.5	16	2 B54877	alpha-conotoxin Pn
34	23.5	17.5	20	2 A34859	heliothermine - Mex
35	23.5	17.5	23	2 F39855	paralytic peptide
36	23.5	17.5	23	2 G39855	paralytic peptide
37	23	17.2	12	1 JTJG0	tremarogen A-10 -
38	23	17.2	12	2 I40663	bma protein - Clos
39	23	17.2	13	2 G22565	R-phycoerythrin ga
40	23	17.2	13	2 PS0453	36K protein 3124 -
41	23	17.2	13	2 PH0799	T-cell receptor al
42	23	17.2	14	2 B61309	lutropin beta chai
43	23	17.2	15	2 A49155	vasotocin-associat
44	23	17.2	15	2 PH0782	T-cell receptor al
45	23	17.2	16	2 PH1634	Ig H chain V-D-J r

ALIGNMENTS

RESULT 1

A42865
Ca2+/calmodulin-dependent myosin light chain kinase (autophosphorylation sites) - rab
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Feb-1997
C:Accession: A42865
R:Gao, Z.H.; Moomaw, C.R.; Hsu, J.; Slaughter, C.A.; Stull, J.T.
Biochemistry 31, 6126-6133, 1992
A:Title: Autophosphorylation of skeletal muscle myosin light chain kinase.
A:Reference number: A42865; MUID:92329432
A:Accession: A42865
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <GAO>
A:Experimental source: skeletal muscle
A:Note: sequence extracted from NCBI backbone (NCBIP:109204)
C:Keywords: calmodulin binding

Query Match 23.1%; Score 31; DB 2; Length 20;
Best Local Similarity 37.5%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 APCGTFSTNTSTDICR 18
||| : |||
Db 2 APGQADQAKAQGDTCR 17

RESULT 2

B49048
T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (fragmen
C:Species: Homo sapiens (man)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
R:Sloud, M.; Kjeldsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.
Eur. J. Immunol. 22, 2413-2418, 1992
A:Title: Limited heterogeneity of T cell receptor variable region gene usage in juven
A:Reference number: A49048; MUID:92387250
A:Accession: B49048
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-18 <STO>
A:Experimental source: patient EV, IL-2R+ synovial T-cells
A:Note: sequence extracted from NCBI backbone (NCBIP:113264)
C:Keywords: T-cell receptor

Query Match 22.4%; Score 30; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 6e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CAPGTFSTNT 11

Db 7 CAPGXYGYT 16
||||:|

RESULT 3

S47381

T-cell antigen receptor VJ junction beta chain - human

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999

C:Accession: S47381

R:Lehner, P.J

Submitted to the EMBL Data Library, August 1994

A:Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T c

A:Reference number: S47355

A:Accession: S47381

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-13 <LEH>

A:Cross-references: EMBL:235698; NID:9527487; PIDN:CAAB4767.1; PID:9527488

C:Keywords: T-cell receptor

Query Match 21.6%; Score 29; DB 2; Length 13;

Best Local Similarity 66.7%; Pred. No. 6.1e+02;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CAPGTFSTNT 10

||| ||||

Db 1 CASSTRSNT 9

RESULT 4

S38292

30K allergen - rye (fragment)

C:Species: Secale cereale (rye)

C>Date: 19-May-1994 #sequence_revision 27-Feb-1997 #text_change 07-May-1999

C:Accession: S38292

R:Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.

A:Title: Comparison of four grass pollen species concerning their allergens of grass gro

A:Reference number: S38288; MUID:94092339

A:Accession: S38292

A:Molecule type: protein

A:Residues: 1-16 <PET>

Query Match 21.6%; Score 29; DB 2; Length 16;

Best Local Similarity 83.3%; Pred. No. 7.3e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PCAPGT 6

|||||

Db 7 PAAPGT 12

RESULT 5

A60317

glucagon-like peptide 1 - marbled electric ray (fragment)

C:Species: Torpedo marmorata (marbled electric ray)

C>Date: 28-Oct-1992 #sequence_revision 28-Oct-1992 #text_change 21-Nov-1997

C:Accession: A60317

R:Conlon, J.M.; Hansen, H.F.; Schwartz, T.W.

Regul. Pept. 13, 94, 1986

A:Title: A truncated glucagon-like peptide I from torpedó pancreas.

A:Reference number: A60317

A:Accession: A60317

A:Molecule type: protein

A:Residues: 1-17 <CON>

C:Superfamily: Glucagon

C:Keywords: duplication; pancreas

Query Match 21.6%; Score 29; DB 2; Length 17;

Best Local Similarity 38.5%; Pred. No. 7.7e+02;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 APGTFSTNTSTD 15

||||:|

Db 2 AEGTYSOVSSUB 14

RESULT 6

I54351

gene HEXA protein - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000

C:Accession: I54351

R:Akili, S.; Chomel, J.C.; Lacorte, J.M.; Bachner, L.; Poenaru, L.

Hum. Mol. Genet. 2, 61-67, 1993

A:Title: Ten novel mutations in the HEXA gene in non-Jewish Tay-Sachs patients.

A:Reference number: I54351; MUID:93258352

A:Accession: I54351

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-21 <RES>

A:Cross-references: GB:S61298; NID:9300412; PIDN:AAD13927.1; PID:94261627

C:Genetics:

A:Gene: GDB:HEXA

A:Cross-references: GDB:120040; OMIM:272800

A:Map position: 15q23-15q24

C:Superfamily: beta-hexosaminidase

Query Match 21.6%; Score 29; DB 2; Length 21;

Best Local Similarity 52.6%; Pred. No. 9.4e+02;

Matches 10; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Qy 5 GT-FSNTTSTSDICR-PHQ 21

||| | | | | |

Db 1 GTFFINKTEIEDFPRPHQ 19

RESULT 7

A39269

IX-1 tumor antigen - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 05-Jan-1996

C:Accession: A39269

R:Rosenbaum, L.C.; Newwelt, E.A.; Van Tol, H.H.M.; Loh, Y.P.; Verbalis, J.G.; Hellstr

Proc. Natl. Acad. Sci. U.S.A. 87, 9928-9932, 1990

A:Title: Expression of neurophysin-related precursor in cell membranes of a small-cel

A:Reference number: A39269; MUID:91088624

A:Accession: A39269

A:Molecule type: protein

A:Residues: 1-22 <ROS>

C:Superfamily: oxytocin-neurophysin

Query Match

21.6%; Score 29; DB 2; Length 22;

Best Local Similarity 80.0%; Pred. No. 9.8e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PCAPG 5

||| |

Db 12 PCGPG 16

RESULT 8

S60565

homeodomain protein hrox3 - California red abalone (fragment)

C:Species: Haliotis rufescens (California red abalone)

C>Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 15-Oct-1999

C:Accession: S60565

R:Deegan, B.M.; Morse, D.E.

Mol. Marine Biol. Biotechnol. 2, 1-9, 1993

A:Title: Identification of eight homeobox-containing transcripts expressed during lar

A:Reference number: S60564; MUID:93372986

A:Accession: S60565

A:Molecule type: mRNA

A:Residues: 1-23 <DEG>

A:Cross-references: EMBL:X79372; NID:g495110; PIDN:CAA55917.1; PID:g495111

C:Genetics:

A:Gene: hrox3

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 20.9%; Score 28; DB 2; Length 23;

Best Local Similarity 57.1%; Pred. No. 1.4e+03;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 16 ICRPHQI 22

:||| :|

Db 5 LCRPRRI 11

RESULT 9

T48881

Leader peptide [Imported] - Vibrio sp.

C:Species: Vibrio sp.

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T48881

R:Xu, Y.; Zhang, Y.; Liang, Z.Y.; Van de Castele, M.; Legrain, C.; Glansdorff, N.

Microbiology 144, 1435-1441, 1998

A:Title: Aspartate carbamoyltransferase from a psychrophilic deep-sea bacterium, *Vibrio*

A:Reference number: 224845

A:Accession: T48881

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-20 <XUY>

A:Cross-references: EMBL:Y09786; PIDN:CAA70922.1

A:Experimental source: strain 2693

Query Match 20.1%; Score 27; DB 2; Length 20;

Best Local Similarity 41.7%; Pred. No. 1.7e+03;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 9 NTTSTSDICRP 20

::|| :|||

Db 7 SSLSSFKLVLRPH 18

RESULT 10

MXKNI

mu-conotoxin GIIIA [validated] - cone shell (*Conus geographus*)

N:Alternate names: geographutoxin I (GTx I); myotoxin I

C:Species: *Conus geographus* (geography cone)

C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 15-Sep-2000

C:Accession: A01786; A23579

R:Sato, S.; Nakamura, H.; Ohizumi, Y.; Kobayashi, J.; Hirata, Y.

FEBS Lett. 155, 277-280, 1983

A:Title: The amino acid sequences of homologous hydroxyproline-containing myotoxins from

A:Reference number: A91309; MUID:83210170

A:Accession: A01786

A:Molecule type: protein

A:Residues: 1-22 <SAV>

R:Cruz, L.J.; Gray, W.R.; Olivera, B.M.; Zeikus, R.D.; Kerr, L.; Yoshikami, D.; Moczydlowski, J.

J. Biol. Chem. 260, 9280-9288, 1985

A:Title: *Conus geographus* toxins that discriminate between neuronal and muscle sodium ch

A:Reference number: A23579; MUID:85261316

A:Accession: A23579

A:Molecule type: protein

A:Residues: 1-22 <CRU>

R:Kohda, D.; Lancelin, J.M.; Inagaki, F.; Wakamatsu, K.

submitted to the Brookhaven Protein Data Bank, December 1992

A:Reference number: A51994; PDB:1TCG

A:Contents: annotation; conformation by (1)H-NMR, residues 1-22

R:Lancelin, J.M.; Kohda, D.; Tate, S.I.; Yanagawa, Y.; Abe, T.; Satake, M.; Inagaki, F.

Biochemistry 30, 6908-6916, 1991

A:Title: Tertiary structure of conotoxin GIIIA in aqueous solution.

A:Reference number: A44639; MUID:91299744

A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR

R:Ott, K.H.; Becker, S.; Gordon, R.D.; Rueterjans, H.

FEBS Lett. 278, 160-166, 1991

A:Title: Solution structure of mu-conotoxin GIIIA analysed by 2D-NMR and distance geo

A:Reference number: A58581; MUID:91122275

A:Contents: annotation; conformation by (1)H-NMR

R:Wakamatsu, K.; Kohda, D.; Hatanaka, H.; Lancelin, J.M.; Ishida, Y.; Oya, M.; Nakamu

Biochemistry 31, 12577-12584, 1992

A:Title: Structure-activity relationships of mu-conotoxin GIIIA: structure determinat

A:Reference number: A44244; MUID:93112598

A:Contents: annotation; conformation by (1)H-NMR

C:Superfamily: mu-conotoxin

C:Keywords: amidated carboxyl end; hydroxyproline; myotoxin; sodium channel inhibitor

F:3-15,4-20,10-21/Disulfide bonds: #status experimental

F:6,7/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

F:17/Modified site: 4-hydroxyproline (Pro) #status experimental

F:22/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 20.1%; Score 27; DB 1; Length 22;

Best Local Similarity 42.9%; Pred. No. 1.8e+03;

Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 17 CRPHQIC 23

:||| :

Db 15 CKPQRC 21

RESULT 11

JN0263

antigen (clone PV12) - Plasmodium vivax (fragment)

C:Species: Plasmodium vivax

C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 09-Sep-1997

C:Accession: JN0263; S21344

R:Ray, P.; Sharma, Y.D.

Biochem. Biophys. Res. Commun. 184, 668-672, 1992

A:Title: Molecular cloning and serological characterization of a new Plasmodium vivax

A:Reference number: JN0263; MUID:92246949

A:Accession: JN0263

A:Molecule type: DNA

A:Residues: 1-15 <RAY>

A:Cross-references: GB:X53681; NID:g10084; PID:g10085

Query Match 19.4%; Score 26; DB 2; Length 15;

Best Local Similarity 54.5%; Pred. No. 1.8e+03;

Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PGTFSTSTST 14

||| :||

Db 3 PQOHSETLVST 13

RESULT 12

B60278

24K antigen - Mycobacterium bovis (fragment)

C:Species: Mycobacterium bovis

C>Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 18-Jun-1993

C:Accession: B60278

R:Fifis, T.; Costopoulos, C.; Radford, A.J.; Bacic, A.; Wood, P.R.

Infect. Immun. 59, 800-807, 1991

A:Title: Purification and characterization of major antigens from a Mycobacterium bo

A:Reference number: A60278; MUID:91147217

A:Accession: B60278

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-16 <PIF>

Query Match 19.4%; Score 26; DB 2; Length 16;

Best Local Similarity 38.5%; Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 APGTFSTSTSTD 15
|||
Db 1 APKYKEELKGT 13

RESULT 13
I54283
arylsulfatase A - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
C:Accession: I54283
R;Regis, S.; Carozzo, R.; Filocamo, M.; Serra, G.; Mastropalo, C.; Gatti, R.
Hum. Genet. 96, 233-235, 1995
A:Title: An AT-deletion causing a frameshift in the arylsulfatase A gene of a late infant
A:Reference number: I54283; MUID:95362256
A:Accession: I54283
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-20 <RES>
A:Cross-references: GB:S78735; NID:gl037139; PIDN:AAB35013.1; PID:gl037140

Query Match 19.4%; Score 26; DB 2; Length 20;
Best Local Similarity 44.4%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 TSSTDICRP 19
::: |||
Db 1 SAHSDHCRP 9

RESULT 14
A59048
convulsant peptide - cone shell (Conus textile)
C:Species: Conus textile (cloth-of-gold cone)
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
C:Accession: A59048
R;Cruz, L.J.; Ramilo, C.A.; Corpuz, G.P.; Olivera, B.M.
Biol. Bull. 183, 159-164, 1992
A:Title: Conus peptides: phylogenetic range of biological activity.
A:Reference number: A59048
A:Accession: A59048
A:Molecule type: protein
A:Residues: 1-23 <CRU>
C:Keywords: amidated carboxyl end; neurotoxin; venom
F;23/Modified site: amidated carboxyl end (Pro) #status predicted

Query Match 19.4%; Score 26; DB 2; Length 23;
Best Local Similarity 27.8%; Pred. No. 2.6e+03;
Matches 5; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 2 CAPGTFSTSTSTDICRP 19
|||
Db 9 CCPPAYCEASG-----CRP 22

RESULT 15
I37144
aspartylglycosaminuria - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I37144
R;Park, H.; Vettese, M.B.; Fensom, A.H.; Fisher, K.J.; Aronson, N.N.
Biochem. J. 290, 735-741, 1993
A:Title: Characterization of three alleles causing aspartylglycosaminuria: two from a Br
A:Reference number: I37144; MUID:93207523
A:Accession: I37144
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA

A:Residues: 1-22 <RES>

A:Cross-references: EMBL:X73071; NID:gl12227; PIDN:CAA51529.1; PID:gl12228

C:Genetics:

A:Gene: AGU

Query Match 19.0%; Score 25.5; DB 2; Length 22;
Best Local Similarity 38.9%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 4; Indels 7; Gaps 1;

QY 6 TFSNTTSTSDICRPHOIC 23

||| |

Db 5 TFSRRVS-----HHIC 15

Search completed: July 13, 2001, 17:17:28
Job time: 102 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:17:11 ; Search time 9.94 Seconds

(without alignments)

79.263 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_185

Perfect score: 134

Sequence: 1 PCAPGTFSTSTSDICRPHQIC 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 1257

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	20.1	22	1	CXMI_CONGE
2	27	20.1	23	1	APP2_BRANA
3	25.5	19.0	23	1	CP23_SPOER
4	25.5	19.0	23	1	PAP2_SPOEX
5	25.5	19.0	23	1	PAP3_SPOEX
6	25	18.7	11	1	TKNA_ONCMY
7	25	18.7	22	1	CXM2_CONGE
8	25	18.7	22	1	CXM3_CONGE
9	24.5	18.3	22	1	LANM_STRMU
10	24	17.9	19	1	HBB2_UROHA
11	23.5	17.5	15	1	DCMM_PSECA
12	23.5	17.5	16	1	CXAB_CONPE
13	23.5	17.5	20	1	HELT_HELHO
14	23.5	17.5	23	1	PAP1_HELVI
15	23.5	17.5	23	1	PAP2_HELVI
16	23	17.2	12	1	TALO_TREME
17	23	17.2	15	1	CIQA_RAT
18	22.5	16.8	23	1	PAP2_MANSE
19	22	16.4	20	1	JHBP_BOMMO
20	22	16.4	23	1	PAP1_SPOEX
21	21.5	16.0	16	1	CXAA_CONPE
22	21	15.7	9	1	R11_SALTY
23	21	15.7	15	1	NUO3_SOLTU
24	21	15.7	18	1	CXAL_CONER
25	21	15.7	20	1	PSBH_SYNVU
26	21	15.7	22	1	SETB_SALTY
27	21	15.7	22	1	TX12_TRIWA
28	20	14.9	19	1	MFH_TRISP
29	20	14.9	20	1	UCRO_EQUAR
30	20	14.9	21	1	MCT3_MOUSE
31	20	14.9	23	1	RL5_HALHA
32	19.5	14.6	15	1	DCMM_PSECH
33	19.5	14.6	23	1	PAP1_MANSE

34	19	14.2	13	1	CXAL_CONST	P15471	conus stria
35	19	14.2	13	1	LMAL_LOCMI	P38496	locusta mig
36	19	14.2	15	1	CXAL_CONGE	P01519	conus geogr
37	19	14.2	15	1	HS11_PINPS	P81083	pinus pinas
38	19	14.2	15	1	RKGG_CARCR	P21586	carretta car
39	19	14.2	15	1	UC08_WAIZE	P80614	zea mays (m
40	19	14.2	15	1	URE2_MORMO	P17338	morganella
41	19	14.2	17	1	ITHB_HIRME	P28502	hirudo medi
42	19	14.2	19	1	CXA2_CONST	P28879	conus stria
43	19	14.2	19	1	LPRM_STAAU	P03063	staphylococ
44	19	14.2	20	1	SODF_PASPI	P81527	pasteurella
45	19	14.2	21	1	FEDB_AMEYE	P80706	amycolatops

ALIGNMENTS

RESULT 1	
CXMI_CONGE	
ID	CXMI_CONGE STANDARD; PRT; 22 AA.
AC	P01523;
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	01-FEB-1995 (Rel. 31, Last annotation update)
DE	MU-CONOTOXIN GLIITA (MYOTOXIN I) (GEOGRAPHUTOXIN I) (GTX-I).
OS	Conus geographus (Geography cone).
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC	Neogastropoda; Conoidea; Conidae; Conus.
OX	NCBI_TaxID=6491;
RN	[1]
RP	SEQUENCE.
RX	MEDLINE=85261316; PubMed=2410412;
RA	Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,
RA	Yoshikami D., Mozdykowski E.;
RT	"Conus geographus toxins that discriminate between neuronal and
RT	muscle sodium channels.";
RL	J. Biol. Chem. 260:9280-9288(1985).
RN	[2]
RP	SEQUENCE.
RX	MEDLINE=83210170; PubMed=6852238;
RA	Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;
RT	"The amino acid sequences of homologous hydroxyproline-containing
RT	myoeloxins from the marine snail Conus geographus venom.";
RL	FEBS Lett. 153:277-280(1983).
RN	[3]
RP	DISULFIDE BONDS.
RX	MEDLINE=90249506; PubMed=2338142;
RA	Hidaka Y., Sato K., Nakamura H., Kobayashi J., Ohizumi Y.,
RA	Simonishi Y.;
RT	"Disulfide pairings in geographutoxin I, a peptide neurotoxin from
RT	Conus geographus.";
RL	FEBS Lett. 264:29-32(1990).
RN	[4]
RP	REVIEW.
RX	MEDLINE=89024586; PubMed=3052286;
RA	Gray W.R., Olivera B.M., Cruz L.J.;
RT	"Peptide toxins from venomous Conus snails.";
RL	Annu. Rev. Biochem. 57:665-700(1988).
RN	[5]
RP	STRUCTURE BY NMR.
RX	MEDLINE=9112275; PubMed=1991506;
RA	Ott K.-H., Becker S., Gordon R.D., Rueterjans H.;
RT	"Solution structure of mu-conotoxin GLIITA analysed by 2D-NMR and
RT	distance geometry calculations.";
RL	FEBS Lett. 278:160-166(1991).
RN	[6]
RP	STRUCTURE BY NMR.
RX	MEDLINE=9129974; PubMed=2069951;
RA	Lancelin J.-M., Kohda D., Tate S.-I., Yanagawa Y., Abe T., Satake M.,
RA	Inagaki F.;
RT	"Tertiary structure of conotoxin GLIITA in aqueous solution.";
RT	Biochemistry 30:6908-6916(1991).
RL	-!- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK

CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM CHANNELS.

DR PIR; A01786; MXKN1.
 DR PIR; A23579; A23579.
 DR PDB; 1TGG; 31-JAN-94.
 DR PDB; 1TCH; 31-JAN-94.
 DR PDB; 1TCJ; 31-JAN-94.
 DR PDB; 1TCK; 31-JAN-94.
 KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;
 3D-structure.
 FT DISULFID 3 15
 FT DISULFID 4 20
 FT DISULFID 10 21
 FT MOD_RES 6 6
 FT MOD_RES 7 7
 FT MOD_RES 17 17
 FT MOD_RES 22 22
 FT MOD_RES 13 16
 FT HELIX 13 21
 FT TURN 19 21
 SQ SEQUENCE 22 AA; 2568 MW; F6CB02ADB359813C CRC64;

Query Match 20.1%; Score 27; DB 1; Length 22;
 Best Local Similarity 42.9%; Pred. No. 5e+02;
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 17 CRPHQIC 23

|||||

Db 15 CKPORCC 21

RESULT 2

APP2_BRANA STANDARD; PRT; 23 AA.

AC P30226;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE CYSTEINE-RICH ANTIFUNGAL PROTEIN 2 (APP2) (FRAGMENT).
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=93138130; PubMed=8422949;
 RA Terras F.R.G., Torrekens S., van Leuven F., Osborn R.W.,
 RA Vanderleyden J., Cammue B.P.A., Broekaert W.F.;
 RT "A new family of basic cysteine-rich plant antifungal proteins from
 Brassicaceae species.";
 RL FEBS Lett. 316:233-240(1993).
 CC -!- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY SENSITIVE TO INORGANIC
 CATIONS.

CC -!- SUBUNIT: FORMS OLIGOMERS IN ITS NATIVE STATE.
 CC -!- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.

DR PIR; S28992; S28992.
 DR HSSP; P30231; IAYJ.
 DR InterPro; IPR002118;
 DR PROSITE; PS00940; GAMMA_THIONIN; PARTIAL.
 KW Fungicide.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2467 MW; A11D4A9E364F5735 CRC64;

Query Match 20.1%; Score 27; DB 1; Length 23;
 Best Local Similarity 28.6%; Pred. No. 5.2e+02;
 Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 GTFSTNTSTDICR 18

|||||

Db 9 GTWSGVCNNACK 22

RESULT 3

CP23_SPOER STANDARD; PRT; 23 AA.
 AC P56683;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CARDIOACTIVE PEPTIDE CAP23.
 OS Spodoptera eridania (Southern armyworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
 OX NCBI_TaxID=37547;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=99196260; PubMed=10098624;
 RA Furuya K., Hackett M., Cirelli M.A., Schegg K.M., Wang H.,
 RA Shabanowitz J., Hunt D.F., Schooley D.A.;
 RT "A cardioactive peptide from the southern armyworm, Spodoptera
 eridania.";
 RL Peptides 20:53-61(1999).
 CC -!- FUNCTION: HAS EXCITATORY EFFECTS ON A SEMI-ISOLATED HEART FROM
 LARVAL MANDUCA SEXTA, CAUSING AN INOTROPIC EFFECT AT LOW
 CONCENTRATIONS OF PEPTIDE AND CHRONOTROPIC AND INOTROPIC EFFECTS
 AT HIGH DOSES.
 CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA; 2519 MW; 0A96D72A70855AE0 CRC64;

Query Match 19.0%; Score 25.5; DB 1; Length 23;
 Best Local Similarity 33.3%; Pred. No. 8.5e+02;
 Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 2 CAPGTFSTNTSTDICRP 19

|||||

Db 7 CTGP-YQTADGR-CKP 21

RESULT 4

PAP2_SPOEX STANDARD; PRT; 23 AA.
 AC P30256;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARALYTIC PEPTIDE II (PP II).
 OS Spodoptera exigua (Beet armyworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
 OX NCBI_TaxID=7107;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=91302298; PubMed=2071576;
 RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Quistad G.B.;
 RT "Isolation and identification of paralytic peptides from hemolymph of
 the lepidopteran insects Manduca sexta, Spodoptera exigua, and
 Heliothis virescens.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 KW Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA; 2477 MW; 0A96CB4600855AE0 CRC64;

Query Match 19.0%; Score 25.5; DB 1; Length 23;
 Best Local Similarity 33.3%; Pred. No. 8.5e+02;
 Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 2 CAPGTFSTSTSDICRP 19
 | | | : | : | : |
 Db 7 CTPG-YQRTADGR--CKP 21

RESULT 5
 ID PAF3_SPOEX STANDARD; PRT; 23 AA.
 AC P30257;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARALYTIC PEPTIDE III (PP III).
 OS Spodoptera exigua (Beet armyworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
 OX NCBI_TaxID=7107;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=91302298; PubMed=2071576;
 RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Oulstad G.B.;
 RT "Isolation and identification of paralytic peptides from hemolymph of
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
 RT Heliothis virescens.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 DR PIR; E39855; E39855.
 KW Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA; 2505 MW; 0A96CB5EB7D55AE0 CRC64;

Query Match 19.0%; Score 25.5; DB 1; Length 23;
 Best Local Similarity 33.3%; Pred. No. 8.5e+02;
 Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 2 CAPGTFSTSTSDICRP 19
 | | | : | : | : |
 Db 7 CTPG-YQRTADGR--CKP 21

RESULT 6
 ID TKNA_ONCMY STANDARD; PRT; 11 AA.
 AC P28499;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SUBSTANCE P.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=92298992; PubMed=1376687;
 RA Jensen J., Conlon J.M.;
 RT "Substance P-related and neurokinin-A-related peptides from the brain
 RT of the cod and trout.";

RL Eur. J. Biochem. 206:659-664(1992).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 DR PIR; S23307; S23307.
 DR PIR; S23308; S23308.
 DR InterPro; IPR002040; -.
 DR PROSITE; PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
 FT MOD_RES 11 11 AMIDATION (BY SIMILARITY).
 SQ SEQUENCE 11 AA; 1358 MW; 214860DEC9D6D1F7 CRC64;

Query Match 18.7%; Score 25; DB 1; Length 11;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RPQH 21
 | | | |
 Db 3 RPQH 6

RESULT 7
 ID CXM2_CONGE STANDARD; PRT; 22 AA.
 AC P01524;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MU-CONOTOXIN GIIIB (MYOTOXIN II) (GEOGRAPHUTOXIN II) (GTX-II).
 OS Conus geographus (Geography cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6491;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85261316; PubMed=2410412;
 RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,
 RA Yoshikami D., Moczydlowski E.;
 RT "Conus geographus toxins that discriminate between neuronal and
 RT muscle sodium channels.";
 RL J. Biol. Chem. 260:9280-9288(1985).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=83210170; PubMed=6852238;
 RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;
 RT "The amino acid sequences of homologous hydroxyproline-containing
 RT myotoxins from the marine snail Conus geographus venom.";
 RL FEBS Lett. 155:277-280(1983).
 RN [3]
 RP REVIEW.
 RX MEDLINE=89024586; PubMed=3052286;
 RA Gray W.R., Olivera B.M., Cruz L.J.;
 RT "Peptide toxins from venomous Conus snails.";
 RL Annu. Rev. Biochem. 57:665-700(1988).
 RN [4]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96280640; PubMed=6888418;
 RA Hill J.M., Alewood P.F., Craik D.J.;
 RT "Three-dimensional solution structure of mu-conotoxin GIIIB, a
 RT specific blocker of skeletal muscle sodium channels.";
 RL Biochemistry 35:8824-8835(1996).
 CC -1- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK
 CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM
 CC CHANNELS.
 DR PIR; A01787; MKKN2.
 DR PIR; B23579; B23579.
 DR PDB; 1GIB; 08-NOV-96.
 KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;
 RW 3D-structure.
 FT DISULFID 3 15

FT DISULFID 4 20
 FT DISULFID 10 21
 FT MOD_RES 6 6
 FT MOD_RES 7 7
 FT MOD_RES 17 17
 FT MOD_RES 22 22
 FT AMIDATION
 SQ SEQUENCE 22 AA; 2599 MW; F50402BA93199E01 CRC64;

Query Match 18.7%; Score 25; DB 1; Length 22;
 Best Local Similarity 44.4%; Pred. No. 9.5e+02;
 Matches 4; Conservative 1; Mismatches 4; Indels 0;

QY 15 DICRPHQIC 23
 | | | | |
 Db 2 DCCTPPK 10

RESULT 8
 CXM3_CONGE STANDARD; PRT; 22 AA.
 AC P05482;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE MU-CONOTOXIN GIIC.
 OS Conus geographus (Geography cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6491;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85261316; PubMed=2410412;
 RA Cruz L.J., Gray W.R., Oliveira B.M., Zeikus R.D., Kerr L.,
 RA Yoshikami D., Moczydlowski E.;
 RT "Conus geographus toxins that discriminate between neuronal and
 RT muscle sodium channels.";
 RL J. Biol. Chem. 260:9280-9288(1985).
 RN [2]

RP REVIEW
 RX MEDLINE=89024586; PubMed=3052286;
 RA Gray W.R., Oliveira B.M., Cruz L.J.;
 RT "Peptide toxins from venomous Conus snails.";
 RL Annu. Rev. Biochem. 57:665-700(1988).
 CC -I- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK
 CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM
 CC CHANNELS.

DR PIR; C23579; C23579.
 DR HSP; P01524; IGIB.
 KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom.
 FT DISULFID 3 15
 FT DISULFID 4 20
 FT DISULFID 10 21
 FT MOD_RES 6 6
 FT MOD_RES 7 7
 FT MOD_RES 17 17
 FT MOD_RES 22 22
 FT AMIDATION
 SQ SEQUENCE 22 AA; 2553 MW; F50402BA92A9813C CRC64;

Query Match 18.7%; Score 25; DB 1; Length 22;
 Best Local Similarity 44.4%; Pred. No. 9.5e+02;
 Matches 4; Conservative 1; Mismatches 4; Indels 0;

QY 15 DICRPHQIC 23
 | | | | |
 Db 2 DCCTPPK 10

RESULT 9
 LANM_STRMU STANDARD; PRT; 22 AA.
 ID LANM_STRMU
 AC P80666;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LANTIBIOTIC MUTACIN B-NY266.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=NY266;
 RX MEDLINE=97379322; PubMed=9237644;
 RA Mota-Meira M., Lacroix C., Lapointe G., Lavoie M.C.;
 RT "Purification and structure of mutacin B-NY266: a new lantibiotic
 RT produced by Streptococcus mutans.";
 RL FEBS Lett. 410:275-279(1997).
 CC -I- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
 CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
 CC TRANSMEMBRANE PORES.
 CC -I- MASS SPECTROMETRY: MW=2270.29; MW ERR=0.21; METHOD=ELECTROSPRAY.
 CC -I- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 DR InterPro; IPR001049; -
 DR Pfam; PF02052; Gallidermin; 1.
 DR PRINTS; PR00323; GALLIDERMIN.
 KW Antibiotic; Bacteriocin; Lantibiotic; Plasmid.
 FT MOD_RES 5 5
 FT MOD_RES 8 8
 FT MOD_RES 14 14
 FT MOD_RES 19 19
 FT THIOETH 3 7
 FT THIOETH 8 11
 FT THIOETH 16 21
 FT THIOETH 19 22
 FT THIOETH 22 22
 SQ SEQUENCE 22 AA; 2425 MW; 961C1480401f92CE CRC64;

Query Match 18.3%; Score 24.5; DB 1; Length 22;
 Best Local Similarity 37.5%; Pred. No. 1.1e+03;
 Matches 6; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 2 CAPGTFSTNTSTDIC 17
 | | | | |
 Db 7 CTPGC-AKTGFSNSYC 21

RESULT 10
 HBB2_UROHA STANDARD; PRT; 19 AA.
 AC P18992;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DE HEMOGLOBIN BETA-2 CHAIN (FRAGMENT).
 OS Uromastix hardwickii (Indian spiny-tailed lizard).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
 OC Uromastix.
 OX NCBI_TaxID=40250;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=84029159; PubMed=6628672;
 RA Naqvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
 RA Joernvall H.;
 RT "Characterization of hemoglobin from the lizard Uromastix
 RT hardwickii.";
 RL FEBS Lett. 162:290-295(1983).
 DR PIR; A05305; A05305.
 DR InterPro; IPR000971; -
 DR PROSITE; PS01033; GLOBIN; PARTIAL.
 KW Heme; Oxygen transport; Respiratory protein; Erythrocyte.
 FT NON_TER 1

SQ SEQUENCE 20 AA; 2156 MW; 91D62B36F7B4F940 CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 20;
Best Local Similarity 41.2%; Pred. No. 1.4e+03;
Matches 7; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

OY 1 PCAPGTF-SNTTSSTDI 16
| | | | |
DB 4 PKLPGLMTSNPDQOTEI 20
| | | | |

RESULT 14
ID PAP1_HELVI STANDARD; PRT; 23 AA.
AC P30251;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARALYTIC PEPTIDE I (PP I).
OS Heliothis virescens (Noctuid moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Heliothinae; Heliothis.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
the lepidopteran insects Manduca sexta, Spodoptera exigu, and
Heliothis virescens."
RL J. Biol. Chem. 266:12873-12877(1991).
CC -/- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -/- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR PIR; F39855; F39855.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2524 MW; 2236CB436D655AFA CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 23;
Best Local Similarity 33.3%; Pred. No. 1.6e+03;
Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

OY 2 CAPGTF-SNTTSSTDI 19
| | | | |
DB 7 CIPG-YMRTADGR--CKP 21
| | | | |

RESULT 15
ID PAP2_HELVI STANDARD; PRT; 23 AA.
AC P30252;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARALYTIC PEPTIDE II (PP II).
OS Heliothis virescens (Noctuid moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Heliothinae; Heliothis.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;

RT "Isolation and identification of paralytic peptides from hemolymph of
the lepidopteran insects Manduca sexta, Spodoptera exigu, and
Heliothis virescens."
RL J. Biol. Chem. 266:12873-12877(1991).
CC -/- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
DR PIR; G39855; G39855.
KW Hemolymph.
SQ SEQUENCE 23 AA; 2508 MW; 2236CB5D6C855AFA CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 23;
Best Local Similarity 33.3%; Pred. No. 1.6e+03;
Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

OY 2 CAPGTF-SNTTSSTDI 19
| | | | |
DB 7 CIPG-YMRTADGR--CKP 21
| | | | |

Search completed: July 13, 2001, 17:19:56
Job time: 165 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:16:56 ; Search time 20.9 Seconds
(without alignments)
145.599 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_185

Perfect score: 134

Sequence: 1 PCAPGTFSTSTSDICRPHQIC 23

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 6111

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP_TREMBL_16.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mammal.*
- 8: sp_mhc.*
- 9: sp_organalle.*
- 10: sp_phase.*
- 11: sp_plant.*
- 12: sp_rhodent.*
- 13: sp_unclassified.*
- 14: sp_vertebrate.*
- 15: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	24.6	15	11 Q9QVD6	Q9QVD6 rattus sp.
2	30	22.4	20	14 Q9QV5	Q9QV5 human immun
3	29	21.6	17	14 Q78381	Q78381 human immun
4	29	21.6	18	14 Q9QEX3	Q9QEX3 human immun
5	29	21.6	21	4 Q16017	Q16017 homo sapien
6	28	20.9	17	14 Q78323	Q78323 human immun
7	28	20.9	17	14 Q78378	Q78378 human immun
8	28	20.9	17	14 Q78327	Q78327 human immun
9	28	20.9	17	14 Q78380	Q78380 human immun
10	28	20.9	21	5 Q25086	Q25086 hermania m
11	28	20.9	23	5 Q25134	Q25134 halotis ru
12	27	20.1	17	14 Q78379	Q78379 human immun
13	27	20.1	20	2 P96173	P96173 vibrio sp.
14	27	20.1	20	5 Q9TWR5	Q9TWR5 phonetria
15	27	20.1	22	13 Q91102	Q91102 morone saxa
16	27	20.1	23	14 Q86611	Q86611 human immun
17	26	19.4	15	5 Q26159	Q26159 plasmodium
18	26	19.4	20	6 Q9TRH7	Q9TRH7 canis famil
19	26	19.4	20	10 Q9S885	Q9S885 lupinus alb

20	26	19.4	20	11 Q9QVD6	Q9QVD6 rattus sp.
21	26	19.4	21	3 Q9URT6	Q9URT6 schizosacch
22	26	19.4	22	3 Q9UR51	Q9UR51 filobasidie
23	26	19.4	23	5 P90716	P90716 beroe ovata
24	26	19.4	23	14 Q9QEX5	Q9QEX5 human immun
25	26	19.4	23	14 Q9QEX4	Q9QEX4 human immun
26	25.5	19.0	22	4 Q13726	Q13726 homo sapien
27	25	18.7	13	10 Q9S922	Q9S922 brassica ol
28	25	18.7	13	11 P97944	P97944 mus musculu
29	25	18.7	17	5 Q9TWC6	Q9TWC6 dirofilaria
30	25	18.7	18	4 Q9UGN8	Q9UGN8 homo sapien
31	25	18.7	19	10 Q9S923	Q9S923 brassica ol
32	25	18.7	20	14 Q78486	Q78486 human immun
33	25	18.7	20	14 Q78507	Q78507 human immun
34	25	18.7	20	14 Q78508	Q78508 human immun
35	25	18.7	20	14 Q78509	Q78509 human immun
36	25	18.7	22	3 Q9Y8F8	Q9Y8F8 glomus moss
37	24.5	18.3	15	4 Q93046	Q93046 homo sapien
38	24	17.9	15	2 Q69142	Q69142 streptococc
39	24	17.9	15	10 Q9S8B9	Q9S8B9 lupinus alb
40	24	17.9	16	4 Q9UC48	Q9UC48 homo sapien
41	24	17.9	16	4 Q9NPQ7	Q9NPQ7 homo sapien
42	24	17.9	16	13 Q9PRU6	Q9PRU6 gallus gall
43	24	17.9	17	13 Q9PRU7	Q9PRU7 gallus gall
44	24	17.9	17	14 Q85719	Q85719 reovirus sp
45	24	17.9	17	14 Q78345	Q78345 human immun

ALIGNMENTS

RESULT 1					
Q9QV5					
ID	Q9QV5	PRELIMINARY;	PRT;	15 AA.	
AC	Q9QV5;				
DT	01-MAY-2000 (TREMREL. 13, Created)				
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)				
DE	01-JUN-2000 (TREMREL. 14, Last annotation update)				
DE	OLIGODENDROCYTE-SPECIFIC UDP-GALACTOSE: CERAMIDE GALACTOSYLTRANSFERASE (FRAGMENT).				
OS	Rattus sp.				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10118;				
RN	[1]				
RP	SEQUENCE.				
RX	MEDLINE=96085162; PubMed=8521863;				
RA	Schulte S., Stoffel W.;				
RT	"UDP galactose:ceramide galactosyltransferase and glutamate/aspartate transporter. Copurification, separation and characterization of the two glycoproteins".				
RL	Eur. J. Biochem. 233:947-953(1995).				
SQ	SEQUENCE 15 AA; 1657 MW; 84474749A06BFFCC CRC64;				

Query Match	24.6%;	Score 33;	DB 11;	Length 15;
Best Local Similarity	66.7%;	Pred. No. 1.2e+02;		
Matches	6;	Conservative 1;	Mismatches 2;	Indels 0;
Gaps	0;			

QY	4	PGTFSNTS 12			
Db	2	PGIFXSTS 10			
RESULT 2					
Q78505					
ID	Q78505	PRELIMINARY;	PRT;	20 AA.	
AC	Q78505;				
DT	01-NOV-1996 (TREMREL. 01, Created)				
DT	01-NOV-1996 (TREMREL. 01, Last sequence update)				
DT	01-NOV-1996 (TREMREL. 08, Last annotation update)				
DE	VIRAL SAMPLE FLO5R5D (FLORIDA LOCAL CONTROL 01), PARTIAL ENV CDS, V5 REGION (FRAGMENT).				

```

OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang L.Q., Leigh-Brown A.J.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RX SEQUENCE FROM N.A.
RM MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
RA Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice.";
RL Science 256:1165-1171(1992).
DR EMBL; M92150; AAA44592.1; -.
FT NON_TER 1
FT NON_TER 20
SQ SEQUENCE 20 AA; 2049 MW; F44F963A48755A07 CRC64;

Query Match 22.4%; Score 30; DB 14; Length 20;
Best Local Similarity 40.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 GTFSTTSSTDIICRP 19
DB 2 GNKGTEETEIRP 16
1 : : : : : : : :
2 : : : : : : : :

RESULT 3
Q78381
ID Q78381 PRELIMINARY; PRT; 17 AA.
AC Q78381;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE VIRAL SAMPLE FLBPR5F (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION
DE (FRAGMENT).
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang L.Q., Leigh-Brown A.J.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RX SEQUENCE FROM N.A.
RM MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
RA Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice.";
RL Science 256:1165-1171(1992).
DR EMBL; M92126; AAA44496.1; -.
FT NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1708 MW; 347570D2D12CA370 CRC64;

Query Match 21.6%; Score 29; DB 14; Length 17;
Best Local Similarity 45.5%; Pred. No. 5.8e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 NTSSTDIICRP 19
DB 3 NNTNTEIRP 13
1 : : : : :
2 : : : : :

RESULT 4
Q9QEX3

```

```

ID Q9QEX3 PRELIMINARY; PRT; 18 AA.
AC Q9QEX3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
OS ENV.
GN Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RX SEQUENCE FROM N.A.
RA Lin H.J., Sivak B.B., Hollinger F.B.;
RT "Mutation rate of human immunodeficiency virus type 1 genomic RNA
deduced from long term culture of its biological clones.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF178667; AA04373.1; -.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 2011 MW; E17BAC9DD31D9910 CRC64;

Query Match 21.6%; Score 29; DB 14; Length 18;
Best Local Similarity 42.9%; Pred. No. 6.1e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 CAPGTFSTTSSTD 15
DB 2 CTELVNTTISTE 15
1 : : : : :
2 : : : : :

RESULT 5
Q16017
ID Q16017 PRELIMINARY; PRT; 21 AA.
AC Q16017;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE HEXA PROTEIN (FRAGMENT).
GN HEXA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93258352; PubMed=8490625;
RA Akli S., Chomel J.C., Lacorte J.M., Bachner L., Poenaru A.,
RA Poenaru L.;
RT "Ten novel mutations in the HEXA gene in non-Jewish Tay-Sachs
patients.";
RL Hum. Mol. Genet. 2:61-67(1993).
DR EMBL; S61298; AAD13927.1; -.
DR HSSP; P06865; IQBC.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2494 MW; D4ACE2D1DA24D8EC CRC64;

Query Match 21.6%; Score 29; DB 4; Length 21;
Best Local Similarity 52.6%; Pred. No. 7.1e+02;
Matches 10; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 5 GT-FSNTTSSTDIICR-PHQ 21
DB 1 GTFFINKTEIEDFFRPHQ 19
1 : : : : :
2 : : : : :

RESULT 6
Q78323
ID Q78323 PRELIMINARY; PRT; 17 AA.
AC Q78323;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE VIRAL SAMPLE FLPAR5A (FLORIDA PATIENT A), PARTIAL ENV CDS, V5 REGION
 DE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L.Q., Leigh-Brown A.J.;
 RA Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92271245; PubMed=1589796;
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
 RA Curran J.W., Jaffe H.W.;
 RT "Molecular epidemiology of HIV transmission in a dental practice."
 RL Science 256:1165-1171(1992).
 DR EMBL; M92109; AAA44465.1; -.
 FT NON_TER 1 1
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 1649 MW; 3E857BBFD12CA370 CRC64;

Query Match 20.9%; Score 28; DB 14; Length 17;
 Best Local Similarity 45.5%; Pred. No. 8.3e+02;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTSSTDICRP 19
 I | : | : |
 DB 3 NNTNGTETFRP 13

RESULT 7
 Q78378 PRELIMINARY; PRT; 17 AA.
 AC Q78378;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1998 (TREMBlrel. 01, Last sequence update)
 DE VIRAL SAMPLE FLPAR5A (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION
 DE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L.Q., Leigh-Brown A.J.;
 RA Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92271245; PubMed=1589796;
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
 RA Curran J.W., Jaffe H.W.;
 RT "Molecular epidemiology of HIV transmission in a dental practice."
 RL Science 256:1165-1171(1992).
 DR EMBL; M92123; AAA44493.1; -.
 FT NON_TER 1 1
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 1723 MW; 34757935D12CA370 CRC64;

Query Match 20.9%; Score 28; DB 14; Length 17;
 Best Local Similarity 45.5%; Pred. No. 8.3e+02;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTSSTDICRP 19
 I | : | : |
 DB 3 NNTNGTETFRP 13

RESULT 8
 Q78327 PRELIMINARY; PRT; 17 AA.
 AC Q78327;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1998 (TREMBlrel. 01, Last sequence update)
 DE VIRAL SAMPLE FLPAR5D (FLORIDA PATIENT A), PARTIAL ENV CDS, V5 REGION
 DE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L.Q., Leigh-Brown A.J.;
 RA Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92271245; PubMed=1589796;
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
 RA Curran J.W., Jaffe H.W.;
 RT "Molecular epidemiology of HIV transmission in a dental practice."
 RL Science 256:1165-1171(1992).
 DR EMBL; M92112; AAA44468.1; -.
 FT NON_TER 1 1
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 1651 MW; 34757BBFD12CA370 CRC64;

Query Match 20.9%; Score 28; DB 14; Length 17;
 Best Local Similarity 45.5%; Pred. No. 8.3e+02;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTSSTDICRP 19
 I | : | : |
 DB 3 NNTNGTETFRP 13

RESULT 9
 Q78380 PRELIMINARY; PRT; 17 AA.
 AC Q78380;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1998 (TREMBlrel. 01, Last sequence update)
 DE VIRAL SAMPLE FLPAR5E (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION
 DE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L.Q., Leigh-Brown A.J.;
 RA Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92271245; PubMed=1589796;
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
 RA Curran J.W., Jaffe H.W.;
 RT "Molecular epidemiology of HIV transmission in a dental practice."
 RL Science 256:1165-1171(1992).
 DR EMBL; M92125; AAA44495.1; -.
 FT NON_TER 1 1
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 1651 MW; 34757BBFD12CA370 CRC64;

Query Match 20.9%; Score 28; DB 14; Length 17;
 Best Local Similarity 45.5%; Pred. No. 8.3e+02;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTTSSTDCICRP 19
 | | : | : |
 Db 3 NNTNGTETFRP 13

RESULT 10
 Q25086 PRELIMINARY; PRT; 21 AA.
 AC Q25086;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE CLONE AH0X4 HOMEBOX PROTEIN (FRAGMENT).
 OS Herdmania momus.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Herdmania.
 OX NCBI_TaxID=7733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EMBRYO;
 RA Kenneth C.V.D.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U09939; AAA18629.1; -.
 DR HSSP; P02833; 9ANT.
 DR InterPro; IPR001356; -.
 DR Pfam; PF00046; homeobox; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON_TER 1
 FT NON_TER 21
 SQ SEQUENCE 21 AA; 2650 MW; AB7FF3AF1FA659C3 CRC64;

Query Match 20.9%; Score 28; DB 5; Length 21;
 Best Local Similarity 57.1%; Pred. No. 1e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 16 ICRPHQI 22
 : | | : |
 Db 4 LCRPRRI 10

RESULT 11
 Q25134 PRELIMINARY; PRT; 23 AA.
 AC Q25134;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HROX3 (FRAGMENT).
 GN HROX3.
 OS Haliotis rufescens (California red abalone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
 OC Haliotidae; Haliotis.
 OX NCBI_TaxID=6454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Degnan B.M.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93372986; PubMed=7689904;
 RA Degnan B.M.; Morse D.E.;
 RT "Identification of eight homeobox-containing transcripts expressed during larval development and at metamorphosis in the gastropod mollusc Haliotis rufescens."
 RT Mol. Mar. Biol. Biotechnol. 2:1-9(1993).
 DR EMBL; X79372; CAA55917.1; -.
 DR HSSP; P02833; 9ANT.
 DR InterPro; IPR001356; -.

DR Pfam; PF00046; homeobox; 1.
 FT NON_TER 1
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2793 MW; CC387AE7BDA6C44D CRC64;

Query Match 20.9%; Score 28; DB 5; Length 23;
 Best Local Similarity 57.1%; Pred. No. 1.1e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 16 ICRPHQI 22
 : | | : |
 Db 5 LCRPRRI 11

RESULT 12
 Q78379 PRELIMINARY; PRT; 17 AA.
 ID Q78379;
 AC Q78379;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE VIRAL SAMPLE FLBR5C (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION (FRAGMENT).
 DE Human immunodeficiency virus type 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L.Q.; Leigh-Brown A.J.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92271245; PubMed=1589796;
 RA Ou C.Y.; Ciesielski C.A.; Myers G.; Banda C.I.; Luo C.C.;
 RA Korber B.T.M.; Mullins J.I.; Schochetman G.; Berkman R.L.;
 RA Econou A.N.; Witte J.J.; Furman L.J.; Satten G.A.; MacInnes K.A.;
 RA Curran J.W.; Jaffe H.W.;
 RT "Molecular epidemiology of HIV transmission in a dental practice."
 RL Science 256:1165-1171(1992).
 DR EMBL; M92124; AAA4494.1; -.
 FT NON_TER 1
 FT NON_TER 17
 SQ SEQUENCE 17 AA; 1652 MW; 34757BBFD1240170 CRC64;

Query Match 20.1%; Score 27; DB 14; Length 17;
 Best Local Similarity 45.5%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTTSSTDCICRP 19
 | | : | : |
 Db 3 NDTNGTETFRP 13

RESULT 13
 P96173 PRELIMINARY; PRT; 20 AA.
 ID P96173;
 AC P96173;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE LEADER PEPTIDE.
 OS Vibrio sp. (strain 2693).
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=79682;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2693;
 RX MEDLINE=98274751; PubMed=9611817;
 RA Xu Y.; Zhang Y.; Liang Z.Y.; Van de Castele M.; Legrain C.;
 RA Glansdorff N.;
 RT "Aspartate carbamoyltransferase from a psychrophilic deep-sea

RT bacterium, Vibrio strain 2693; properties of the enzyme, genetic
 RT organization and synthesis in *Escherichia coli*.;
 RL Microbiology 144:1435-1441(1998).
 DR EMBL; Y09786; CAA/0922.1; -.
 SQ SEQUENCE 20 AA; 2241 MW; 35C31F588FBB5D63 CRC64;

Query Match 20.1%; Score 27; DB 2; Length 20;
 Best Local Similarity 41.7%; Pred. No. 1.4e+03;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 9 NTSSTDICRPH 20
 : : : : :
 Db 7 SSSLSEKLVLRPH 18

RESULT 14

Q9TWR5 PRELIMINARY; PRT; 20 AA.
 AC Q9TWR5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE PNV2 TOXIN (FRAGMENT).
 OS Phoneutria nigriventer (Brazilian armed spider).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
 OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
 OX NCBI_TaxID=6918;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94030062; PubMed=8216354;
 RA Bento A.C., Novello J.C., Marangoni S., Antunes E., Giglio J.R.,
 RA Oliveira B., de Nucci G.;
 RT "Identification of a new vascular smooth muscle contracting
 RT polypeptide in Phoneutria nigriventer spider venom.";
 RL Biochem. Pharmacol. 46:1092-1095(1993).
 SQ SEQUENCE 20 AA; 2176 MW; F28C3D81D983BCA5 CRC64;

Query Match 20.1%; Score 27; DB 5; Length 20;
 Best Local Similarity 80.0%; Pred. No. 1.4e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 15 DICRP 19
 : : : : :
 Db 6 DICQP 10

RESULT 15

Q91102 PRELIMINARY; PRT; 22 AA.
 AC Q91102;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HOX-B3-LIKE HOMEODOMAIN PROTEIN (FRAGMENT).
 OS Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RX MEDLINE=95005122; PubMed=7921046;
 RA Pavell A.M., Stellwag E.J.;
 RT "Survey of Hox-like genes in the teleost Morone saxatilis:
 RT implications for evolution of the Hox gene family.";
 RL Mol. Mar. Biol. Biotechnol. 3:149-157(1994).
 DR EMBL; U09944; AAC59650.1; -.
 DR HSSP; P02833; 9ANT.
 DR InterPro; IPR001356; -.

DR Pfam; PF00046; homeobox; 1.
 KW Homeobox; Nuclear protein; DNA-binding.
 FT NON_TER 1 1
 FT NON_TER 22 22
 SQ SEQUENCE 22 AA; 2703 MW; AE4485CB7FF7CF1D CRC64;

Query Match 20.1%; Score 27; DB 13; Length 22;
 Best Local Similarity 42.9%; Pred. No. 1.5e+03;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 16 ICRPHOI 22
 : : : : :
 Db 2 LCRPRRV 8

Search completed: July 13, 2001, 17:19:39
 Job time: 163 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:14:30 ; Search time 17.98 Seconds
(without alignments)
77.550 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_185
Perfect score: 134
Sequence: 1 PCAPGTFNWTSTDICRPHQIC 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 169437

Minimum DB seq length: 0
Maximum DB seq length: 23

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	34.3	17	20	AAW95323
2	41	30.6	17	21	AAV51961
3	41	30.6	17	21	AAV51973
4	37	27.6	22	21	AAV51542
5	37	27.6	23	20	AAW73416
6	34.5	25.7	23	19	AAW65481
7	33.5	25.0	20	21	AAV23019
8	33	24.6	15	19	AAV20893
9	33	24.6	16	16	AAV75932
10	32.5	24.3	23	6	AAV50661
11	32.5	24.3	23	20	AAV28912

12	32	23.9	15	18	AAW07656	Human ATM gene pro
13	32	23.9	15	18	AAW06235	ATM epitope w1. H
14	32	23.9	15	19	AAW77164	Pharmaceutically a
15	32	23.9	15	20	AAV33126	Human umbilical co
16	32	23.9	16	20	AAV14391	Peptide #19 for ep
17	32	23.9	16	20	AAV15781	Antigenic peptide
18	32	23.9	17	21	AAV59437	Human delta3 fragm
19	32	23.9	20	21	AAV65710	TGF beta 2 mutant
20	32	23.9	22	18	AAV26538	Erythropoietin rec
21	31	23.1	15	20	AAV08939	Fibronectin-like s
22	31	23.1	15	21	AAV52487	Fibronectin-deri
23	31	23.1	15	21	AAV15600	Immunopeptide deri
24	31	23.1	20	16	AAV84512	Hepatitis C virus
25	31	23.1	20	17	AAV91006	HCV E2 peptide E2
26	30	22.4	12	12	AAV10375	N-terminal sequenc
27	30	22.4	15	13	AAV21621	Sequence encoded b
28	30	22.4	18	15	AAV48347	Conserved sequence
29	30	22.4	19	18	AAV13656	Erythropoietin rec
30	30	22.4	19	21	AAV17318	EPO-mimetic peptid
31	30	22.4	20	19	AAV83054	Human Fas peptide
32	30	22.4	20	21	AAV14762	Human Fas epitope,
33	30	22.4	20	21	AAV90912	Human Fas extracel
34	30	22.4	20	21	AAV69518	Human Fas peptide
35	29.5	22.0	16	19	AAV37752	Antigenic C-termin
36	29.5	22.0	17	6	AAV50866	Hepatitis B virus
37	29.5	22.0	23	6	AAV50862	Sequence of immuno
38	29.5	22.0	23	19	AAV21337	Human semaphorin I
39	29	21.6	6	17	AAV11030	Immunomodulatory p
40	29	21.6	6	17	AAV11024	Immunomodulatory p
41	29	21.6	6	18	AAV44953	Immunomodulatory p
42	29	21.6	6	19	AAV65277	Cysteine analogue-
43	29	21.6	6	19	AAV52036	Monomer having a t
44	29	21.6	6	19	AAV51947	Peptide having imm
45	29	21.6	6	20	AAV09379	Immunoactive pepti

ALIGNMENTS

RESULT 1
AAW95323
ID AAW95323 standard; Protein; 17 AA.
XX AC AAW95323;
XX DT 15-MAR-1999 (first entry)
XX DE Costant and variable domain sequence of C. psittaci CPS92-106.
XX KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;
KW antiporphyrin acid; immune response; infection; diagnostic; assay; MOMP;
KW major outer membrane protein; autoimmune; inflammatory; porphyria;
KW Ebstein Barr virus; antioxidant.
XX OS Chlamydia psittaci.
XX PN WO9850074-A2.
XX PD 12-NOV-1998.
XX PF 06-MAY-1998; 98WO-US09237.
XX PR 18-FEB-1998; 98US-0025521.
PR 06-MAY-1997; 97US-0045689.
PR 06-MAY-1997; 97US-0045739.
PR 06-MAY-1997; 97US-0045779.
PR 06-MAY-1997; 97US-0045780.
PR 06-MAY-1997; 97US-0045784.
PR 06-MAY-1997; 97US-0045787.
PR 14-AUG-1997; 97US-0911593.
PR 18-FEB-1998; 98US-0025174.
XX 18-FEB-1998; 98US-0025176.

PA (UYVA-) UNIV VANDERBILT.
 XX Mitchell WM, Stratton CW;
 PI WPI; 1999-059653/05.
 XX
 XX Composition with two agents effective against different stages of
 PT chlamydial life cycle : comprises agent targetted against cryptic
 PT phase, against elementary body phase, against replicating phase,
 PT probenicid and antiporphyric
 XX
 XX Claim 4; Fig 3; 138pp; English.
 XX
 CC The invention relates to the diagnosis and management of infections by
 CC Chlamydia species. The invention provides a composition that comprises
 CC at least two agents, where each of the agents is effective against a
 CC different phase of the chlamydial life cycle. The agents are selected
 CC from: (a) agents targetted against cryptic phase of chlamydial life
 CC cycle; (b) agents targetted against elementary body phase of chlamydial
 CC life cycle; (c) agents targetted against replicating phase of chlamydial
 CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
 CC is used to elicit a protective immune response to Chlamydia infection in
 CC an animal or human and is applied until the animal or human tests
 CC negative for Chlamydia infection. It is also used to treat biological
 CC material infected with Chlamydia. Diagnostic kits for antibody assays
 CC against recombinant major outer membrane protein (MOMP), and for DNA
 CC amplification assays for chlamydial genes, are used to diagnose disease,
 CC e.g. autoimmune disease, an inflammatory disease or a disease that
 CC occurs in an immuno-compromised individual, associated with Chlamydia
 CC infection. The kits are used to detect chlamydial elementary bodies in a
 CC sample. They are also used to monitor and/or modify the course of therapy
 CC in a patient. The treatment reduces the acellular load of infectious
 CC Ebsstein Barr virus. The method is also used to treat porphyria, by
 CC reducing the number of elementary bodies and applying a drug, e.g.
 CC cimetidine, and antioxidants, to reduce the adverse effects associated
 CC with porphyria. Sequences AA95320 to AA95323 represent constant and
 CC variable domain sequences of various Chlamydia species.
 XX
 XX Sequence 17 AA;
 SQ

Query Match 34.3%; Score 46; DB 20; Length 17;
 Best Local Similarity 64.3%; Pred. No. 3.9;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CAPGTFSTSTSD 15
 ||||| : 1
 Db 1 casgtsnttvaad 14

RESULT 2
 AAY51961
 ID AAY51961 standard; peptide; 17 AA.
 XX
 XX AAY51961;
 XX
 XX 23-JUN-2000 (first entry)
 DT
 XX P. pastoris lysyl oxidase peptide fragment #11.
 DE
 XX Lysyl oxidase; lipoxxygenase; protein disulfide isomerase; phenol oxidase;
 KW peroxidase; protein disulfide reductase; tyrosine oxidase; fodder;
 KW sulfhydryl oxidase; food additives.
 XX
 XX Pichia pastoris.
 OS
 XX DE19840069-Al.
 PN
 XX 09-MAR-2000.
 PD
 XX 03-SEP-1998; 98DE-1040069.
 PF
 XX 03-SEP-1998; 98DE-1040069.
 PR

XX (BADI) BASF AG.
 XX Friedrich T, Bewert W, Lueddecke E, Klingler J, Heger R;
 PI WPI; 2000-257743/23.
 XX
 XX Manufacture of active preparations comprises cross linking a protein,
 PT which surrounds the active substance with an enzyme, especially a novel
 PT lysyl oxidase from Pichia pastoris
 XX
 XX Claim 17; Page 17; 22pp; German.
 PS
 XX This invention describes a novel method to manufacture a preparation of
 CC an active substance, where the active substance is surrounded by at
 CC least one layer consisting of a protein that is cross-linked by an
 CC enzyme chosen from the group of lipoxxygenase, protein disulfide
 CC isomerase, phenol oxidase and peroxidase, lysyl oxidase, protein
 CC disulfide reductase, tyrosine oxidase or sulfhydryl oxidases. Enzymes
 CC chosen from lipoxxygenase, protein disulfide isomerase, phenol oxidase
 CC and peroxidase, protein disulfide reductase, tyrosine oxidase or
 CC sulfhydryl oxidases, especially lysyl oxidase are useful for formulation
 CC of preparations of active substances. The method of the invention is
 CC used for manufacturing preparations of active substances. The
 CC preparations are useful as food additives or fodder or as
 CC pharmaceuticals. AAY51951-Y51962 represent fragments of the Pichia
 CC pastoris lysyl oxidase which are used to illustrate the method of the
 CC invention.
 XX
 XX Sequence 17 AA;
 SQ

Query Match 30.6%; Score 41; DB 21; Length 17;
 Best Local Similarity 70.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNT 10
 ||||| : 1
 Db 7 pcapgvvynt 16

RESULT 3
 AAY51973
 ID AAY51973 standard; peptide; 17 AA.
 XX
 XX AAY51973;
 XX
 XX 23-JUN-2000 (first entry)
 DT
 XX P. pastoris lysyl oxidase fragment #11.
 DE
 XX Lysyl oxidase; lipoxxygenase; protein disulfide isomerase; phenol oxidase;
 KW peroxidase; protein disulfide reductase; tyrosine oxidase; food;
 KW sulfhydryl oxidase; animal feed.
 XX
 XX Pichia pastoris.
 OS
 XX DE19840489-Al.
 PN
 XX 09-MAR-2000.
 PD
 XX 04-SEP-1998; 98DE-1040489.
 PF
 XX 04-SEP-1998; 98DE-1040489.
 PR
 XX (BADI) BASF AG.
 XX Friedrich T, Bewert W, Lueddecke E, Klingler J, Heger R;
 PI WPI; 2000-272257/24.
 XX
 XX Use of specified enzymes, especially lysyl oxidase, as protein
 PT crosslinking agents for formulating compositions containing active

PT ingrédiants -
XX Claim 17; Page 16; 20pp; German.
PS
XX This invention describes a novel method where an enzyme (I) selected
CC from lipoxigenases, protein disulfide isomerases, phenol oxidases and
CC peroxidases, lysyl oxidases, protein disulfide reductases, tyrosine
CC oxidases or sulphydryl oxidases is used to formulate compositions
CC containing active ingredients. (I) is useful for crosslinking protein
CC layers surrounding active ingredients in food, animal feed and
CC pharmaceutical products. The compositions can be formulated without
CC using chemical crosslinking agents. AAY51963-V51974 represent fragments
CC of the Pichia pastoris lysyl oxidase protein which is used to
XX illustrate the method of the invention.
SQ Sequence 17 AA;
Query Match 30.6%; Score 41; DB 21; Length 17;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 PCAPGTFSTNT 10
DB 7 pcapgyvnynt 16
IIII II
PCAPGTFSTNT 10
IIII II
PCAPGYVNYNT 16
RESULT 4
AAB51542
ID AAB51542 standard; Peptide; 22 AA.
AC AAB51542;
XX
XX 15-FEB-2001 (first entry)
XX Yada homologous peptide #5.
XX
XX Proteobacteria; extracellular domain; virulence determinant; Yada;
KW adhesin; proteobacterial infection prevention; vaccine.
XX
XX Thiobacillus ferrooxidans.
XX
XX WO200061165-A1.
XX
XX 19-OCT-2000.
XX
XX 13-APR-2000; 2000WO-US09866.
XX
XX 13-APR-1999; 99US-0129073.
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Lupas AN;
XX
XX WPI; 2000-647397/62.
XX
XX An isolated polypeptide conserved in proteobacterial extracellular
PT domains used in the treatment and prevention of bacterial infections -
PT
XX Example 5; Page 59; 85pp; English.
PS
XX This invention relates to peptides AAB51512 - AAB51537 which represent
CC conserved proteobacterial extracellular domains. Sequences
CC AAB51538 - AAB51618 represent peptides homologous to Yada, a yersinia
CC adhesin which is an important virulence determinant of the Yersinia
CC species. The invention includes an antibody which binds to the
CC proteobacterial extracellular peptides, and an immunogenic composition
CC containing the antibody used as a vaccine to prevent infection by a
CC proteobacteria. The polypeptides and antibodies are useful in the
CC treatment and prevention of proteobacterial infections. The polypeptides
CC can also be used to identify compounds which antagonize the binding of a
CC bacterial adhesion to its ligand. The host cell can be used to produce

CC the polypeptides in a suitable culture system. The composition can be
XX used to vaccinate a patient against a proteobacterial infection.
SQ Sequence 22 AA;
Query Match 27.6%; Score 37; DB 21; Length 22;
Best Local Similarity 50.0%; Pred. No. 94;
Matches 10; Conservative 1; Mismatches 7; Indels 2; Gaps 1;
QY 3 APGTFSTNTSSDTCRPHQI 22
DB 5 apgtlsqt--stdavngsqi 22
IIII II II II
APGTFSTNTSSDTCRPHQI 22
IIII II II II
APGTLSQT--STDVNGSQL 22
RESULT 5
AAW73416
ID AAW73416 standard; Protein; 23 AA.
XX
AC AAW73416;
XX
XX 19-FEB-1999 (first entry)
XX
XX Human secreted protein encoded by Gene No. 20.
XX
XX Secreted protein; human; protein therapy; gene therapy; blood disorder;
KW pathological condition; diagnosis; cancer; neurological disorder;
KW developmental abnormality; foetal deficiency; leukaemia; hepatic disease;
KW immune system disorder; Alzheimer's disease; cognitive disorder;
KW schizophrenia; prostate disease; autoimmune disorder; AIDS.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Misc-difference 23 /note= "unspecified amino acid"
FT
XX WO9854206-A1.
PN
XX 03-DEC-1998.
PD
XX
XX 28-MAY-1998; 98WO-US10868.
PF
XX 29-AUG-1997; 97US-0056296.
PR 30-MAY-1997; 97US-0044039.
PR 30-MAY-1997; 97US-0048093.
PR 30-MAY-1997; 97US-0048101.
PR 30-MAY-1997; 97US-0048190.
PR 30-MAY-1997; 97US-0048356.
PR 30-MAY-1997; 97US-0050935.
PR 29-AUG-1997; 97US-0056250.
PR 29-AUG-1997; 97US-0056293.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Carter KC, Dillon PJ, Endress GA, Feng P, NI J;
PI Rosen CA, Ruben SM, Yu G;
XX
XX WPI; 1999-070209/06.
DR N-PSDB; AAV08830.
XX
XX New isolated human genes - useful for diagnosis and treatment of,
PT e.g. cancers, neurological disorders, immune diseases, developmental
PT disorders or blood disorders
XX
XX Claim 11; Page 157; 188pp; English.
XX
XX This sequence is encoded by a cDNA of the invention, designated
CC Gene No. 20. This sequence represents a human secreted protein, and is
CC expressed ubiquitously, including T-cells and amygdala.
CC The DNA sequences of the invention and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions

CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the DNA sequences.
 CC Specific uses are described for each of the DNA sequences and the encoded
 CC proteins, based on which tissues they are most highly expressed in, and
 CC include developing products for the diagnosis or treatment of cancer,
 CC tumours, neurological disorders, developmental abnormalities and foetal
 CC deficiencies, blood disorders, leukaemias, diseases of the immune system
 CC (including allergies or asthma), hepatic disease, Alzheimer's and
 CC cognitive disorders, schizophrenia, prostate diseases, autoimmune
 CC disorders and AIDS. The polypeptides are also useful for identifying
 CC their binding partners.

XX Sequence 23 AA;

Query Match 27.6%; Score 37; DB 20; Length 23;

Best Local Similarity 43.88; Pred. No. 99;

Matches 7; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAPGTFNSTSTSDIC 17

Db 5 cgpaaagtacssacac 20

RESULT 6

AAW65481
 ID AAW65481 standard; peptide; 23 AA.

XX AC AAW65481;

XX DT 12-OCT-1998 (first entry)

XX Hepatitis B surface antigen derived peptide (IGP 1082).

XX Annexin V; hepatitis B surface antigen; immunogen; vaccine;

XX hepatitis delta virus; infection; HBSAg.

XX Synthetic.

OS Hepatitis b virus.

XX WO9829442-A1.

XX 09-JUL-1998.

XX 23-DEC-1997; 97WO-EP07268.

XX 11-JUL-1997; 97EP-0870103.

XX 30-DEC-1996; 96EP-0870164.

XX (INNO-) INNOGENETICS NV.

XX De Meyer S, Depla E, Maertens G, Yap S;

XX WPI; 1998-388040/33.

XX Immunogenic polypeptide from hepatitis B surface antigen - useful
 in, e.g. vaccine against hepatitis B virus or hepatitis delta virus
 infection

XX Example 3; Page 35; 71pp; English.

XX The invention relates to an immunogenic peptide derived from hepatitis B
 CC surface antigen (HBSAg) which competes with the hepatitis B surface
 CC antigen/annexin V interaction or which binds a compound or antibody
 CC competing with the hepatitis B surface antigen/annexin V interaction.
 CC Also claimed are: (1) a combination of the immunogenic peptide and a
 CC negatively charged phospholipid; (2) a peptide composition comprising
 CC the immunogenic peptide; (3) a vaccine comprising the immunogenic peptide
 CC as an active substance; (4) antibodies which specifically bind to the
 CC peptide and inhibit binding of HBSAg to annexin V, and (5) a therapeutic
 CC composition comprising as an active substance the antibodies of (4).
 CC The vaccine of (3), and the therapeutic composition of (5), can be used
 CC as an immunoculum to vaccinate humans against an infection with hepatitis

CC B and/or hepatitis delta virus. The immunogenic peptide can be used in a
 CC method to detect antibodies which are capable of competing with the
 CC hepatitis B and/or hepatitis delta virus surface antigen/annexin V
 CC interaction. The immunogenic peptide can also be used to screen for
 CC drugs which block the binding between annexin V and the peptide, and as
 CC a therapeutic to treat humans infected with hepatitis B virus and/or
 CC hepatitis delta virus. The present sequence represents one of the
 CC peptide fragments derived from HBSAg which were synthesised to map the
 CC annexin V-binding site on HBSAg.

XX Sequence 23 AA;

Query Match 25.7%; Score 34.5; DB 19; Length 23;

Best Local Similarity 44.4%; Pred. No. 2.2e+02;

Matches 8; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 1 PCAPGTFNSTSTSDICR 18

Db 1 pllpgt---ststgpc 15

RESULT 7

AAB23019
 ID AAB23019 standard; peptide; 20 AA.

XX AC AAB23019;

XX 16-JAN-2001 (first entry)

XX Human APC protein 20 aa repeat #4 (1643-1662).

XX APC gene; Adenomatous Polyposis Coli gene; human; chromosome 5q21;

XX familial adenomatous polyposis; FAP locus; Gardner's syndrome; GS;

XX sporadic tumour; adenoma; carcinoma; cancer; lung; breast; colon; rectum;

XX bladder; liver; sarcoma; stomach; prostate; leukaemia; lymphoma;

XX tumour suppressor; anti-APC antibody; detection; diagnosis; prognosis;

XX genetic predisposition; drug screening; DP2.5; repeat region.

XX Homo sapiens.

XX US6114124-A.

XX 05-SEP-2000.

XX 25-MAY-1995; 95US-0450582.

XX 16-JAN-1991; 91GB-0000962.

XX 16-JAN-1991; 91GB-0000963.

XX 16-JAN-1991; 91GB-0000974.

XX 16-JAN-1991; 91GB-0000975.

XX 08-AUG-1991; 91US-0741940.

XX 12-AUG-1994; 94US-0289548.

XX (ICIL) IMPERIAL CHEM IND PLC.

XX (UJJO) UNIV JOHNS HOPKINS.

XX (UTAH) UNIV UTAH.

XX (CANC-) CANCER INST.

XX Carlson M, Groden J, Joslyn G, Kinzler K, Markham AF, Anand R;

XX Albertsen H, White RL, Thliveris A, Nakamura Y, Vogelstein B;

XX Hedge PJ;

XX WPI; 2000-565003/52.

XX Detecting Adenomatous Polyposis Coli (APC) protein in a sample for
 CC diagnosing cancers, involves contacting the sample with antibodies that
 CC specifically bind to APC protein and detecting the complex formed -
 XX Example 15; Column 33-34; 125pp; English.

XX The invention relates to a novel method for detecting Adenomatous
 CC Polyposis Coli (APC) protein in a sample. The method involves

CC contacting the sample with antibodies which specifically binds to the
 CC 2843 amino acid form of the human APC protein, or to a mutant APC
 CC protein, and detecting an APC-antibody complex. Mutations in the APC
 CC gene play a role in tumorigenesis, indicating that it is a tumour
 CC suppressor gene. It is located on chromosome 5q21, which corresponds to
 CC the FAP (familial adenomatous polyposis) locus. FAP is an autosomal
 CC dominant inherited disease in which affected individuals develop
 CC hundreds to thousands of adenomatous polyps in the colon and rectum,
 CC some of which progress to malignancy. The FAP locus is often found to
 CC be deleted in sporadic (i.e., non-familial) adenomas and carcinomas, and
 CC chromosome 5q deletions have also been observed in tumours of the lung,
 CC breast, colon, rectum, bladder, liver, sarcomas, stomach, and prostate,
 CC and in leukaemias and lymphomas. Although the FAP locus contains
 CC several other genes such as FBR, TBL1, TBL2, and MCC, it is thought that
 CC mutations in the APC gene play a key role in the development of FAP and
 CC sporadic tumours. The method is useful for detecting APC protein and its
 CC mutant forms in foetal tissue, placental tissue, amniotic fluid, blood,
 CC serum or a tumour sample. The method is useful for diagnosing or
 CC prognosing neoplastic tissue, for detecting a genetic predisposition to
 CC cancer, for detecting germline and somatic alteration of wild-type APC
 CC genes, and for testing therapeutic agents for the ability to suppress
 CC tumours. Sequences AAB23016-B23022 represent seven 20 amino acid repeats
 CC that are semiregularly spaced in the human APC protein.

XX Sequence 20 AA;

Query Match 25.0%; Score 33.5; DB 21; Length 20;
 Best Local Similarity 44.4%; Pred. No. 2.7e+02;
 Matches 8; Conservative 2; Mismatches 5; Indels 3; Gaps 1;

QY 2 CAPGT---FSNWTSTSDI 16

DB 2 cvegtplnfstatslsdl 19

RESULT 8

AAAY20893
 ID AAY20893 standard; Protein; 15 AA.

AC AAY20893;

DT 22-JUL-1999 (first entry)

DE Human presenilin I mutant protein fragment 39.

Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 frameshift mutation; age-related disease; neurodegenerative disorder;
 Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 Huntington's disease; multiple sclerosis; alcoholic liver disease;
 diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.

OS Homo sapiens.

PN WO9845322-A2.

PD 15-OCT-1998.

PF 02-APR-1998; 98WO-IB00705.

PR 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

PA (UYRO-) UNIV ROTTERDAM ERASMUS.

PI Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX
 DR

WPI: 1998-609901/51.
 N-PSDB: AAX75761.

Diagnosing disease by detecting frameshift mutations in RNA or
 corresponding protein mutations - used to diagnose cancer and
 neurological diseases, particularly Alzheimer's disease, and also
 for treatment and prevention with specific ribozymes or wild-type
 RNA

Disclosure; Figure 10; 258pp; English.

This invention describes a novel method for the diagnosis of a disease
 caused by, or associated with, an RNA molecule that has a frameshift
 mutation. The method is used to diagnose age-related diseases, especially
 cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 and many others listed) or susceptibility to these disorders. The method
 allows a definitive diagnosis of Alzheimer's disease in living patients,
 at an early stage. It is based on the observation that disease may be
 caused by mutations in RNA rather than DNA. The invention describes the
 use of neuronal system RNA molecules, specifically proteins including
 beta-amyloid precursor protein (beta-APP), the microtubule associated
 proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 protein-C (HMGP-C) and neuroendocrine specific protein A.

Sequence 15 AA;

Query Match 24.6%; Score 33; DB 19; Length 15;

Best Local Similarity 83.3%; Pred. No. 2.3e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAPGTG 7

DB 4 ccpgtf 9

RESULT 9

AAAR75932

ID AAR75932 standard; Protein; 16 AA.

AC AAR75932;

DT 24-FEB-1996 (first entry)

DE Dextranase N-terminal sequence.

XX Dextranase; thermostable enzyme; dextran hydrolysis; sugar.

OS Penicillium minioluteum HI-4.

PN EP663443-A1.

PD 19-JUL-1995.

PF 13-DEC-1994; 94EP-0203614.

PR 14-DEC-1993; 93CU-0000115.

XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.

XX Delgado Boada JM, Campana HR, Curbelo DM, Jimenez ER;

PI Patron CF, Cremata Alvarez JA, Garcia Fernandez R;

PI Garcia Garcia BM, Gonzalez Martinez ME, Herrera Martinez LS;

PI Margollez Clark E, Morera Cordova V, Raices Perez-Castaneda MR;

XX WPI: 1995-247530/33.

KW phosphatidylinositol-3 kinase; PI-3; cancer predisposition; carrier;
 KW untranslated region; UTR.
 XX
 OS Homo sapiens.
 XX WO9636695-A1.
 XX
 XX 21-NOV-1996.
 XX
 XX 16-MAY-1996; 96WO-US07040.
 XX
 XX 28-JUL-1995; 95US-0508836.
 XX 16-MAY-1995; 95US-0441822.
 XX 21-JUN-1995; 95US-0493092.
 XX
 XX (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Collins FS, Shiloh Y, Tagle DA;
 XX
 XX WPI; 1997-012074/01.
 DR
 XX
 XX New gene ATM implicated in ataxia-telangiectasia and related protein
 PT - useful in screening methods, partic. for identifying disease
 PT carriers
 XX
 XX Claim 24; Page 86; 127pp; English.
 PS
 XX AAW07656-62 are human ATM protein epitopes used to generate polyclonal
 CC and monoclonal antibodies. Antibodies raised against the ATM protein
 CC detected mono-specifically a high molecular weight of the expected size
 CC of 350 kDa on Western blots of protein lysates derived from fibroblast
 CC and lympho- blastoid cell lines. Because of the high frequency of
 CC truncation mutations in the ATM gene, mutated ATM protein can be
 CC identified if such proteins are stable. Mutations in the ATM gene cause
 CC ataxia- telangiectasia (A-T), a progressive genetic disorder affecting
 CC the central nervous and immune systems. The ATM gene, located at
 CC chromosome 11q22-23, is probably involved in a novel signal transduction
 CC system that links DNA damage surveillance to cell cycle control. The ATM
 CC gene product (AAW07655) has a highly conserved C-terminal region showing
 CC high sequence homology to the catalytic domain of phosphatidylinositol-3
 CC kinases. A-T mutations affect a variety of tissues and lead to cancer
 CC predisposition. Identification of A-T carriers, by analysis at nucleic
 CC acid or protein levels, allows better supervision and treatment of such
 CC subjects who are at increased risk of developing cancer and are
 CC particularly sensitive to radiation. The transgenic animals and
 CC transformed cells are useful as models of the human disease. Also viral
 CC vectors expressing the ATM protein can be used in gene therapy of A-T.
 XX
 XX Sequence 15 AA;
 SQ
 Query Match 23.9%; Score 32; DB 18; Length 15;
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 8 SNTTSSSTDIC 17
 |: : |||:|
 Db 6 ssasqstdic 15
 RESULT 13
 AAW06235
 ID AAW06235 standard; peptide; 15 AA.
 XX
 AC AAW06235;
 XX
 XX 13-AUG-1997 (first entry)
 DT
 XX ATM epitope #1.
 DE
 XX ATM gene; ataxia-telangiectasia; progressive genetic disorder; antibody;
 KW central nervous system; immune system; chromosomal instability; therapy;
 KW

KW cancer predisposition; radiation sensitivity; cell cycle abnormality;
 KW multi-system disease; autosomal recessive; cerebellar ataxia; cerebellum;
 KW general motor dysfunction; Purkinje cell; oculocutaneous telangiectasia;
 KW blood vessel; bulbar conjunctiva; facial skin; A-T.
 XX
 OS Homo sapiens.
 XX WO9636691-A1.
 XX
 XX 21-NOV-1996.
 XX
 XX 16-MAY-1996; 96WO-US07025.
 XX
 XX 08-APR-1996; 96US-0629001.
 XX 16-MAY-1995; 95US-0441822.
 XX
 XX (KOHN/) KOHN K I.
 PA (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
 XX
 XX Shiloh Y;
 PI
 XX
 XX WPI; 1997-012070/01.
 DR
 XX
 XX New isolated ataxia-telangiectasia gene - used to develop prods. for
 PT the study, diagnosis and treatment of ataxia-telangiectasia.
 PT
 XX
 XX Disclosure; Page 16; 153pp; English.
 PS
 XX AAW06235-W06238 represent epitopes of the ATM protein (see AAW06234 for
 CC full length protein) which are specifically recognised by antibodies
 CC targeted against ATM. Ataxia-telangiectasia (A-T) is caused by
 CC mutations, insertions, or deletions in the coding region of the ATM gene
 CC (see AAT43497 for wild type open reading frame). A-T is a progressive
 CC genetic disorder affecting the central nervous and immune systems. A-T
 CC involves chromosomal instability, cancer predisposition, radiation
 CC sensitivity, and cell cycle abnormalities. A-T is a multi-system disease
 CC inherited in an autosomal recessive manner. Cerebellar ataxia that
 CC gradually develops into general motor dysfunction is the first clinical
 CC hallmark of A-T, and results from progressive loss of Purkinje cells in
 CC the cerebellum. Oculocutaneous telangiectasia (dilation of blood vessels)
 CC develops in the bulbar conjunctiva and facial skin, and is later
 CC accompanied by graying of the hair and atrophic changes in the skin. The
 CC co-occurrence of cerebellar ataxia and telangiectases in the conjunctivae
 CC and occasionally on the facial skin (the second early hallmark of A-T)
 CC usually establishes the differential diagnosis of A-T from other
 CC cerebellar ataxias. The gene, can be used in methods for detecting
 CC carriers of a defective gene that causes A-T. The methods and antibodies
 CC can be used in the study, diagnosis and therapy of A-T.
 XX
 XX Sequence 15 AA;
 SQ
 Query Match 23.9%; Score 32; DB 18; Length 15;
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 8 SNTTSSSTDIC 17
 |: : |||:|
 Db 6 ssasqstdic 15
 RESULT 14
 AAW77164
 ID AAW77164 standard; peptide; 15 AA.
 XX
 AC AAW77164;
 XX
 XX 23-NOV-1998 (first entry)
 DT
 XX Pharmaceutically active peptide 5.
 DE
 XX Pharmaceutically active peptide; target; organ; lymphocyte; treatment;
 KW pharmaceutical agent; disease; radioactive isotope; imaging agent.
 KW

```

XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9839469-A1.
XX PD 11-SEP-1998.
XX PF 04-MAR-1998; 98WO-US04188.
XX PR 04-MAR-1997; 97US-0810074.
XX PR 04-MAR-1997; 97US-0039509.
XX PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX PI Belkind A, Golan I, Hagai Y, Lazarovits J, Levanon A;
XX PI Nimrod A, Panet A, Vogel T, Zeelon E;
XX DR WPI; 1998-495863/42.
XX PT New peptide(s) binding targets in organs and lymphocytes - for the
XX PT targetted delivery of toxins, anti-cancer drugs and cardiovascular
XX PT agents to arteries, veins, placenta, liver
XX PS Claim 10; Page 92; 114pp; English.
XX SS Sequences shown in AAW77160 to AAW77214 and AAW79167 represent
XX CC non-naturally- occurring pharmacuetically active peptides. These novel
XX CC peptides specifically bind to undetermined and determined targets in
XX CC various organs and in lymphocytes. The peptides can be used in
XX CC compositions, where they can be linked to pharmaceutical agents, to treat
XX CC various diseases and conditions. The peptides or chimeric polypeptides
XX CC comprising these pharmacuetically active peptides and a second peptide
XX CC may be labelled with a marker (radioactive isotope, etc) to form an
XX CC imaging agent. This agent is used to bind an organ so that the organ can
XX CC be imaged.
XX CC
XX CC Sequence 15 AA;
XX SQ

```

Query Match 23.9%; Score 32; DB 19; Length 15;
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

Qy 8 SNTTSSTDICRPHQ 21
Db 2 sttrnrdinkptq 15

```

RESULT 15
 AAY33126
 ID AAY33126 standard; Protein; 15 AA.
 XX AC AAY33126;
 XX DT
 XX DT 16-NOV-1999 (first entry)
 XX DE Human umbilical cord specific epitope from clone TUV-R4B*-#23.
 XX KW Tissue specific epitope; umbilical cord; human; endothelial cell; vein;
 XX KW tumor cell; tumor tissue; delivery agent; imaging; artery; radial; liver;
 XX KW coronary; mammary; safenal; femoral; placenta; kidney; heart;
 XX KW central nervous system.
 XX OS Homo sapiens.
 XX PN WO9945020-A1.
 XX PD 10-SEP-1999.
 XX PF 04-MAR-1999; 99WO-US04691.
 XX PR 04-MAR-1998; 98WO-US04188.

```

PR 10-SEP-1998; 98US-0154404.
XX PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX PI Vogel T, Panet A;
XX DR WPI; 1999-540813/45.
XX PT New peptides useful for directing therapeutics to cancer tissues - and
XX PT endothelial cells -
XX PS Example 1; Page 67; 70pp; English.
XX CC This invention describes novel non-naturally occurring pharmacuetically
XX CC active peptide epitopes (I) comprising the tripeptide, Glu-Gly-Arg. (I)
XX CC are useful for binding to endothelial cell and tumor cells/tissues, and
XX CC are useful as delivery agents. The products described in the invention
XX CC are also useful for imaging an organ, preferably an artery (umbilical
XX CC cord artery, radial artery, coronary artery or mammary artery), a vein
XX CC (umbilical cord vein, safenal vein or femoral vein), placenta, tumor
XX CC tissue, kidney, heart or liver. The artery is preferably damaged, and is
XX CC preferably an artery (umbilical cord artery, radial artery, coronary
XX CC artery or a mammary artery, preferably a damaged coronary artery), vein
XX CC (umbilical cord vein, safenal vein or femoral vein), placenta, tumor
XX CC tissue, kidney, heart, liver, or central nervous system. AAY33123-Y33132
XX CC represent epitopes described in the method of the invention.
XX CC
XX CC Sequence 15 AA;
XX SQ

```

Query Match 23.9%; Score 32; DB 20; Length 15;
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

Qy 8 SNTTSSTDICRPHQ 21
Db 2 sttrnrdinkptq 15

```

Search completed: July 13, 2001, 17:16:50
 Job time: 140 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:14:50 ; Search time 12.12 Seconds
(without alignments)
38.228 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_185

Perfect score: 134

Sequence: 1 PCAPTFTSNTSSTDCRPHQIC 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 110791

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	53.7	15	1	US-08-221-583-45
2	72	53.7	15	1	US-08-221-583-46
3	72	53.7	15	5	PCT-US95-04018-45
4	72	53.7	15	5	PCT-US95-04018-46
5	71	53.0	15	1	US-08-221-583-44
6	71	53.0	15	5	PCT-US95-04018-44
7	57	42.5	15	1	US-08-221-583-47
8	57	42.5	15	5	PCT-US95-04018-47
9	48	35.8	20	2	US-08-126-016-24
10	44	32.8	15	1	US-08-221-583-48
11	44	32.8	15	5	PCT-US95-04018-48
12	39	29.1	23	4	US-09-101-146-58
13	33	24.6	16	1	US-08-354-618-3
14	33	24.6	20	2	US-08-934-915-141
15	32	23.9	15	1	US-08-493-092-4
16	32	23.9	15	1	US-08-508-836A-4
17	32	23.9	15	2	US-08-629-001A-4
18	32	23.9	15	4	US-08-642-274D-4
19	32	23.9	15	4	US-08-952-127-4
20	32	23.9	22	1	US-08-484-635-178
21	32	23.9	22	2	US-08-484-631-178
22	32	23.9	22	2	US-08-827-570-178
23	31.5	23.5	20	1	US-08-318-193-63
24	31	23.1	18	6	5217891-2
25	31	23.1	20	2	US-08-934-915-78
26	31	23.1	20	4	US-08-612-973-83
27	30	22.4	18	1	US-08-323-531-8

28	30	22.4	18	1	US-08-198-094-8	Sequence 8, Appl1
29	30	22.4	18	4	US-08-107-794A-8	Sequence 8, Appl1
30	30	22.4	18	5	PCT-US93-07424-8	Sequence 8, Appl1
31	30	22.4	18	5	PCT-US95-02087-8	Sequence 8, Appl1
32	30	22.4	19	1	US-08-484-635-244	Sequence 244, App
33	30	22.4	19	2	US-08-484-631-244	Sequence 244, App
34	30	22.4	19	2	US-08-827-570-244	Sequence 244, App
35	29.5	22.0	16	1	US-08-574-763-7	Sequence 7, Appl1
36	29	21.6	6	4	US-08-750-142B-20	Sequence 20, Appl1
37	29	21.6	9	1	US-08-054-860-14	Sequence 14, Appl1
38	29	21.6	9	3	US-08-442-378-14	Sequence 14, Appl1
39	29	21.6	12	4	US-08-602-999A-280	Sequence 280, App
40	29	21.6	15	1	US-08-221-583-49	Sequence 49, Appl1
41	29	21.6	15	4	US-08-602-999A-315	Sequence 315, App
42	29	21.6	15	5	PCT-US95-04018-49	Sequence 49, Appl1
43	29	21.6	22	2	US-08-124-981A-28	Sequence 28, Appl1
44	29	21.6	22	3	US-09-037-190-31	Sequence 31, Appl1
45	29	21.6	22	3	US-09-037-192-31	Sequence 31, Appl1

ALIGNMENTS

RESULT 1
US-08-221-583-45
; Sequence 45, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595r1s
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-45

Query Match 53.7%; Score 72; DB 1; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.00075;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GTFSTNTSSTDCRCP 19

Db 1 GTFSTNTSSTDIARP 15

RESULT 2

```
US-08-221-583-46
; Sequence 46, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-46

Query Match 53.7%; Score 72; DB 1; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.00075;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 SNTTSSTDICRPHQI 22
Db 1 SNTTSSTDICRPHQI 15

RESULT 3
PCT-US95-04018-45
; Sequence 45, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
```

```
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-45

Query Match 53.7%; Score 72; DB 5; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.00075;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GTFSTTSSTDICRP 19
Db 1 GTFSTTSSTDICRP 15

RESULT 4
PCT-US95-04018-46
; Sequence 46, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
```

NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-46

Query Match 53.7%; Score 72; DB 5; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.00075;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 SNTTSSSTDICRPHQI 22
|||||
Db 1 SNTTSSSTDICRPHQI 15

RESULT 5
US-08-221-583-44
; Sequence 44, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25.mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-44

Query Match 53.0%; Score 71; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APGTFSTNTSSTDI 16
|||||
Db 2 APGTFSTNTSSTDI 15

RESULT 6
PCT-US95-04018-44
; Sequence 44, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-44

Query Match 53.0%; Score 71; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APGTFSTNTSSTDI 16
|||||
Db 2 APGTFSTNTSSTDI 15

RESULT 7
US-08-221-583-47
; Sequence 47, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia

; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-04018-47

Query Match 42.5%; Score 57; DB 1; Length 15;
Best Local Similarity 91.7%; Pred. No. 0.074;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQI 22
Db 1 TSSTDICRPHQI 12

RESULT 8
PCT-US95-04018-47
; Sequence 47, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994

; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-04018-47

Query Match 42.5%; Score 57; DB 5; Length 15;
Best Local Similarity 91.7%; Pred. No. 0.074;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQI 22
Db 1 TSSTDICRPHQI 12

RESULT 9
US-08-126-016-24
; Sequence 24, Application US/08126016
; Patent No. 5811261
; GENERAL INFORMATION:
; APPLICANT: WALLACH, DAVID
; APPLICANT: NOPHAR, YARON
; APPLICANT: KEMPER, OLIVER
; APPLICANT: ENGELMANN, HARTMUT
; APPLICANT: BRAKEBUSCH, CORD
; APPLICANT: ADERKA, DAN
; TITLE OF INVENTION: EXPRESSION OF THE RECOMBINANT TUMOR
; NUMBER OF SEQUENCES: 26
; NUMBER OF INVENTIONS: NECROSIS FACTOR BINDING PROTEIN I (TBP-I)
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Nelmark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,016
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/625668
; FILING DATE: 13-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, ROGER L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-126-016-24

Query Match 35.8%; Score 48; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCAPGTF 8
Db 12 PCAPGTF 19

RESULT 10

US-08-221-583-48
; Sequence 48, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221.583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-48

Query Match 32.8%; Score 44; DB 1; Length 15;
Best Local Similarity 88.9%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 TDICRPHQI 22
Db 1 TDIARPHQI 9

RESULT 11

PCT-US95-04018-48
; Sequence 48, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-48

Query Match 32.8%; Score 44; DB 5; Length 15;
Best Local Similarity 88.9%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 TDICRPHQI 22
Db 1 TDIARPHQI 9

RESULT 12

US-09-101-146-58
; Sequence 58, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998

```
;
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-101-146-58

Query Match 29.1%; Score 39; DB 4; Length 23;
Best Local Similarity 43.8%; Pred. No. 29;
Matches 7; Conservative 2; Mismatches 7; Indels 7; Gaps 0;

QY 2 CAPOTFNTSTSDIC 17
   |||||
DB 4 CAAGTTGATGTTATC 19

RESULT 13
US-08-354-618-3
; Sequence 3, Application US/08354618
; Patent No. 5637491
; GENERAL INFORMATION:
; APPLICANT: Campana, Hernan Roca
; APPLICANT: Garcia, Bianca Maria Garcia
; APPLICANT: Clark, Emilio Margollez
; APPLICANT: Curbelo, Dania Mateu
; APPLICANT: Boada, Julio Marcos Delgado
; APPLICANT: Martinez, Luis S. Herrera
; APPLICANT: Alvarez, Jos Alberto Cremata
; APPLICANT: Perez-Casta eda, Manuel Rafael Raices
; APPLICANT: Martinez, Maria Elena Gonz lez
; APPLICANT: Jim nez, Bfrain Rodriguez
; TITLE OF INVENTION: Dextranase enzyme, method for its
; PRODUCTION AND DNA ENCODING THE ENZYME
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ronald J. Baron
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.0 version B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/354, 618
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CU 115/93
; FILING DATE: 14-December-1993
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
```

```
US-08-354-618-3

Query Match 24.6%; Score 33; DB 1; Length 16;
Best Local Similarity 46.2%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 GTFSTNTSTSDIC 17
   |||||
DB 2 GTTNTHCGADFC 14

RESULT 14
US-08-934-915-141
; Sequence 141, Application US/08934915
; Patent No. 5932412
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; APPLICANT: CHENG, HWEI-MING
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN
; PAPILLOMAVIRUS 1, 5, 6, 8,
; TITLE OF INVENTION: 11, 16, 31, 33 AND 56,
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR
; DIAGNOSTIC PURPOSES
; NUMBER OF SEQUENCES: 193
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MASON & ASSOCIATES, P.A.
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
; CITY: CLEARWATER
; STATE: FLORIDA
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 3.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,915
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/949,836
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LOUISE A. Foutch
; REGISTRATION NUMBER: 37,133
; REFERENCE/DOCKET NUMBER: 1946.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 813-538-3800
; TELEFAX: 813-538-3820
; TELEX:
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-934-915-141

Query Match 24.6%; Score 33; DB 2; Length 20;
Best Local Similarity 53.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 8 SNTTSTSDICRPH 20
   |||||
DB 7 SNEVSSPEIRQH 19

RESULT 15
US-08-493-092-4
; Sequence 4, Application US/08493092
```

Patent No. 5728807
GENERAL INFORMATION:
APPLICANT: Shiloh, Yosef
APPLICANT: Tagle, Danilo A.
APPLICANT: Collins, Francis S.
TITLE OF INVENTION: Ataxia-Telangiectasia Gene
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reising, Ethington, Barnard & Perry
STREET: P.O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
ZIP: 48099
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/493,092
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-310 (TAV)
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-493-092-4

Query Match 23.9%; Score 32; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 8 SNTSSSTDIC 17
Db 6 SSASQSTDLC 15

Search completed: July 13, 2001, 17:17:09
Job time: 139 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:12:20 ; Search time 13.32 Seconds
(without alignments)
223.034 Million cell updates/sec

Title: us-09-800-909-2_copy_163_201

Perfect score: 216

Sequence: 1 PCAPGTFSTSTDCRPHQICNVVAIPGNASMDAVCT 39

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 8589

Minimum DB seq length: 0

Maximum DB seq length: 39

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38.5	17.8	37	2 G49050	T-cell surface gly
2	38	17.6	33	2 A37479	huwentoxin-I - Chi
3	34	15.7	37	2 S68261	hypothetical prote
4	33.5	15.5	39	2 A05323	phospholipase A2 (
5	33	15.3	39	2 G82613	hypothetical prote
6	32.5	15.0	35	2 A60959	agelenin - funnel-
7	32	14.8	30	2 A22977	delta-endotoxin -
8	31	14.4	20	2 A42865	Ca2+/calmodulin-de
9	31	14.4	30	2 S70343	napin large chain
10	31	14.4	34	2 E49410	t-complex polypept
11	31	14.4	39	2 A60716	somatotropin intro
12	30.5	14.1	38	2 A48158	pheromone precursor
13	30.5	13.9	18	2 B49048	T-cell receptor be
14	30	13.9	26	2 S28994	antifungal protein
15	30	13.9	29	2 C61233	conceptus protein
16	30	13.9	35	2 A39830	dentin matrix, inc
17	30	13.9	38	2 A45495	beta-defensin-1 -
18	29	13.4	13	2 S47381	T-cell antigen rec
19	29	13.4	16	2 S38292	30K allergen - rye
20	29	13.4	17	2 A00317	glucagon-like pept
21	29	13.4	21	2 I34351	gene HEXA protein
22	29	13.4	22	2 A39269	Lx-1 tumor antigen
23	29	13.4	24	2 S08293	lectin - sunn hemp
24	29	13.4	26	2 B42865	Ca2+/calmodulin-de
25	29	13.4	31	2 S53233	gene X protein - h
26	29	13.4	32	2 I48415	heat shock factor
27	29	13.4	34	2 D81044	hypothetical prote
28	28	13.0	18	1 A58589	alpha-conotoxin EI
29	28	13.0	23	2 S60565	homeodomain protei

30 28 13.0 25 2 S74094
31 28 13.0 27 2 S28995
32 28 13.0 27 2 C44636
33 28 13.0 27 2 T12330
34 28 13.0 29 1 GCFLE
35 28 13.0 29 2 A61135
36 28 13.0 30 1 TIFULW
37 28 13.0 30 2 A47607
38 28 13.0 31 2 S21743
39 28 13.0 32 2 E82089
40 27.5 12.7 30 2 S28991
41 27.5 12.7 31 2 A55430
42 27.5 12.7 35 2 PL0164
43 27.5 12.7 36 2 A82208
44 27.5 12.7 36 2 JT0513
45 27 12.5 20 2 A39328

fibulin 1 variant
antifungal protein
homeotic protein H
metallothionein -
glucagon - Europea
glucagon - bigeye
trypsin inhibitor
immunogenic protei
trypsin inhibitor
hypothetical prote
antifungal protein
conotoxin NgVIA -
alpha-lactalbumin
hypothetical prote
Ig heavy chain V-I
notechis II-5b non

ALIGNMENTS

RESULT 1

G49050

T-cell surface glycoprotein CD8 beta-1 chain, secreted form 5 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-1994 #sequence_revision 03-Nov-1995 #text_change 29-Aug-1997
C:Accession: G49050

R:DiSanto, J.P.; Smith, D.; de Bruin, D.; Lacy, E.; Flomenberg, N.

A:Title: Transcriptional diversity at the duplicated human CD8 beta loci.
A:Reference number: A49050; MUID:93170376

A:Accession: G49050

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-37 <DIS>

A:Note: Sequence extracted from NCBI backbone (NCBIP:125543)

C:Genetics:

A:Gene: GDB:CD8B1; CD8B

A:Cross-references: GDB:119771; OMIM:186730

A:Map position: 2p12-2p12

C:Keywords: alternative splicing; extracellular protein; glycoprotein

Query Match 17.8%; Score 38.5; DB 2; Length 37;
Best Local Similarity 34.8%; Pred. No. 2.1e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 1 PCAPGTFSTSTDCRPHQIC 23

Db 13 PLSPNACMDTTA---ILQPHRSC 32

RESULT 2

A37479

huwentoxin-I - Chinese bird spider

C:Species: Selenocosmia huwena (Chinese bird spider)

C:Date: 18-Mar-1994 #sequence_revision 07-Oct-1994 #text_change 16-Feb-1996

C:Accession: A37479; JCI089

R:Liang, S.P.; Zhang, D.Y.; Pan, X.; Chen, Q.; Zhou, P.A.

Toxicon 31, 969-978, 1993

A:Title: Properties and amino acid sequence of huwentoxin-I, a neurotoxin purified fr

A:Reference number: A37479; MUID:94024948

A:Accession: A37479

A:Molecule type: protein

A:Residues: 1-33 <LIA>

R:Liang, S.P.; Zong, X.; Luo, J.C.; Jing, H.; Gu, X.C.

Acta Sci. Natur. Univ. Pekin. 29, 668-674, 1993

A:Title: Secondary structure study of huwentoxin-I, a neurotoxin from the venom of th

A:Reference number: JCI089

A:Accession: JCI089

A:Molecule type: protein

A:Residues: 1-33 <LI2>

C:Comment: This peptide is the major active protein component of venom in this specie

C;Keywords: presynaptic neurotoxin; venom
F;2-17,9-22,16-29/Disulfide bonds: #status experimental

Query Match 17.6%; Score 38; DB 2; Length 33;
Best Local Similarity 30.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 5 GTFSNTTSSTDICRPHQICN 24
| | | | | : | : | : | :
Db 4 GVFDACTFGRNECCPNVCS 23

RESULT 3

S68261
hypothetical protein gadd7.2 - long-tailed hamster
C;Species: Cricetus longicaudatus (long-tailed hamster)
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S68261
R;Hollander, M.C.; Alamo, I.; Fornace Jr., A.J.
Nucleic Acids Res. 24, 1589-1593, 1996
A;Title: A novel DNA damage-inducible transcript, gadd7, inhibits cell growth, but lacks
A;Reference number: S68260; MUID:96211359
A;Accession: S68261
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-37 <HOL>
A;Cross-references: EMBL:L40430

Query Match 15.7%; Score 34; DB 2; Length 37;
Best Local Similarity 60.0%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 23 CNVVAIPGNA 32
| | | | |
Db 10 CTVTRAFGNA 19

RESULT 4

A05323
Phospholipase A2 (EC 3.1.1.4) - Mexican beaded lizard (fragment)
N;Alternate names: phosphatidylcholine 2-acylhydrolase
C;Species: Heloderma horridum (Mexican beaded lizard)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 05-Apr-1995
C;Accession: A05323
R;Sosa, B.P.; Alegon, A.C.; Martin, B.M.; Possani, L.D.
Biochemistry 25, 2927-2933, 1986
A;Reference number: A05323; MUID:86243292
A;Note: H. h. horridum
A;Accession: A05323
A;Molecule type: protein
A;Residues: 1-39 <SOS>
C;Comment: Phospholipases A2 catalyze the calcium-dependent hydrolysis of the 2-acyl gro
C;Superfamily: phospholipase A2
C;Keywords: calcium; carboxylic ester hydrolase; lipid degradation; venom

Query Match 15.5%; Score 33.5; DB 2; Length 39;
Best Local Similarity 31.0%; Pred. No. 9.6e+02;
Matches 9; Conservative 3; Mismatches 10; Indels 7; Gaps 2;

QY 2 CAPGTFSN-----TTSSTDI-CRPHQIC 23
| | | : | : | : | : | : | : | :
Db 11 CGAGNAASDYSQLGTEKTDTCMRDHDHC 39

RESULT 5

G82613
hypothetical protein XF1988 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: G82613

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717
A;Note: for a complete list of authors see reference number A59328 below

A;Accession: G82613
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <SIM>

A;Cross-references: GB:AE004018; GB:AE003849; NID:9107093; PIDN:NAF84790.1; GSPDB:GN
A;Experimental source: Strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marcks, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1988

Query Match 15.3%; Score 33; DB 2; Length 39;
Best Local Similarity 37.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 6 TFSNTTSSTDICRPHQ 21
: | | | | : | : | : | :
Db 24 SLKVTVSADVMRAHR 39

RESULT 6

A60959
agelenin - funnel-weaving spider (Agelena opulenta)
C;Species: Agelena opulenta
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 24-Feb-1994
C;Accession: A60959
R;Hagiwara, K.; Sakai, T.; Miwa, A.; Kawai, N.; Nakajima, T.
Biomed. Res. 11, 181-186, 1990
A;Title: Complete amino acid sequence of a new type of neurotoxin from the venom of t
A;Reference number: A60959
A;Accession: A60959
A;Molecule type: protein
A;Residues: 1-35 <HAG>
C;Keywords: calcium channel inhibitor; presynaptic neurotoxin; venom

Query Match 15.0%; Score 32.5; DB 2; Length 35;
Best Local Similarity 42.9%; Pred. No. 1.2e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 17 CRPH-QICNVVAIP 29
| | | : | : | : | : | : | : | :
Db 3 CLPHNRCNALSGP 16

RESULT 7

A22977
delta-endotoxin - Bacillus thuringiensis (fragment)
C;Species: Bacillus thuringiensis
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 07-Feb-1997
C;Accession: A22977
R;Armstrong, J.L.; Rohrmann, G.F.; Beaudreau, G.S.
J. Bacteriol. 161, 39-46, 1985
A;Reference number: A22977; MUID:85104736
A;Accession: A22977

A:Molecule type: protein
A:Residues: 1-30 <ARM>
C:Superfamily: 28K parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 14.4%; Score 32; DB 2; Length 30;
Best Local Similarity 24.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 4 PGTFSNTTSSDTCRPHQICNVVAI 28

Db 5 PNEINLLSINEIDNPVILQAI 29

RESULT 8

A42865
Ca2+/calmodulin-dependent myosin light chain kinase (autophosphorylation sites) - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Feb-1997
C:Accession: A42865
R:Gao, Z.H.; Moomaw, C.R.; Hsu, J.; Slaughter, C.A.; Stull, J.T.
Biochemistry 31, 6126-6133, 1992
A:Title: Autophosphorylation of skeletal muscle myosin light chain kinase.
A:Reference number: A42865; MUID:92329432

A:Accession: A42865

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <GAO>

A:Experimental source: skeletal muscle

A>Note: sequence extracted from NCBI backbone (NCBIP:109204)

C:Keywords: calmodulin binding

Query Match 14.4%; Score 31; DB 2; Length 20;
Best Local Similarity 37.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 3 APGTFSNTTSSDTCR 18

Db 2 APGDAQAKAQDTCR 17

RESULT 9

S70343
napin large chain L2B - Swedish turnip (fragments)
C:Species: Brassica napus var. rapifera (Swedish turnip, rutabaga)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 01-Dec-2000
C:Accession: S70343; S70342

R:Neumann, G.M.; Condron, R.; Thomas, I.; Polya, G.M.

Biochim. Biophys. Acta 1295, 34-43, 1996

A:Title: Purification and sequencing of multiple forms of Brassica napus seed napin large

A:Reference number: S70340; MUID:96283791

A:Accession: S70343

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-19; 20-24; 25-30 <NEU>

A:Accession: S70342

A>Status: preliminary

A:Molecule type: protein

A:Residues: 4-19 <NE2>

C:Superfamily: wheat alpha-amylase inhibitor

Query Match 14.4%; Score 31; DB 2; Length 30;
Best Local Similarity 40.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 21 QICNVVAIPG 30

Db 19 RVCNIRTPG 28

RESULT 10

E49410
t-complex polypeptide 1 homolog (peak 2 fraction) - rabbit (fragments)
N:Alternate names: chaperonin homolog (peak 2)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 21-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 04-Sep-1998
C:Accession: E49410
R:Kommelmeier, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.;
Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993
A:Title: Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven r
A:Reference number: A49410; MUID:94089752

A:Accession: E49410

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-34 <ROM>

A:Experimental source: reticulocyte

A>Note: sequence modified after extraction from NCBI backbone

A>Note: sequence extracted from NCBI backbone (NCBIP:141043)

C:Superfamily: molecular chaperone t-complex-type

Query Match 14.4%; Score 31; DB 2; Length 34;

Best Local Similarity 38.9%; Pred. No. 1.8e+03;

Matches 7; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

Qy 19 PHQICNVVAIPGNASMDA 36

Db 17 PRQLCD-----NAGFDA 28

RESULT 11

A60716
somatotropin intron-related protein RDE.25 - rat (fragment)
N:Alternate names: growth hormone gene-related protein RDE.25
C:Species: rattus norvegicus (Norway rat)
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-May-1999
C:Accession: A60716
R:Montpetit, M.L.; Tenniswood, M.P.

J. Cell. Biochem. 39, 285-292, 1989

A:Title: Does the lack of regression-associated mRNA expression render a rat ventral

A:Reference number: A60716; MUID:89214371

A:Accession: A60716

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-39 <MON>

C:Comment: The mRNA encoding this hypothetical protein shows homology at the nucleoti
he complete mRNA may encode a protein related to somatotropin through abnormal splici

Query Match 14.4%; Score 31; DB 2; Length 39;

Best Local Similarity 27.6%; Pred. No. 2e+03;

Matches 8; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 2 CAPGTFSNTTSSDTCRPHQICNVVAIPG 30

Db 10 CLPEVLSSIPSTHVAHHQLQNLVPSSG 38

RESULT 12

A48158
pheromone precursor MF alpha - fungus (Filobasidium floriforme)

C:Species: Filobasidiella neoformans, Cryptococcus neoformans

C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994

C:Accession: A48158

R:Moore, T.D.; Edman, J.C.

Mol. Cell. Biol. 13, 1962-1970, 1993

A:Title: The alpha-mating type locus of Cryptococcus neoformans contains a peptide ph

A:Reference number: A48158; MUID:93180845

A:Accession: A48158

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-38 <MOO>

A>Note: sequence inconsistent with nucleotide translation

A;Note: sequence extracted from NCBI backbone (NCBIN:126064, NCBIPI:126070)

Query Match 14.1%; Score 30.5; DB 2; Length 38;
Best Local Similarity 35.7%; Pred. No. 2.2e+03;
Matches 10; Conservative 4; Mismatches 9; Indels 5; Gaps 1;

OY 7 FSNVTSSTDCRPHQICNVVAIPGNASM 34
|||:::| | | | |
Db 12 FSAATSSSEAPR-----NQEAHPGGMTL 34

RESULT 13

B49048
T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C:Accession: B49048
R:Sloud, M.; Kjeidsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.
Eur. J. Immunol. 22, 2413-2418, 1992
A;Title: Limited heterogeneity of T cell receptor variable region gene usage in juvenile
A;Reference number: A49048; MUID:92387250
A;Accession: B49048
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-18 <SIO>
A;Experimental source: patient EV, IL-2R+ synovial T-cells
A;Note: sequence extracted from NCBI backbone (NCBIP:113264)
C;Keywords: T-cell receptor

Query Match 13.9%; Score 30; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 CAPGTFSNTT 11
|||: |
Db 7 CAPGXYYGYT 16

RESULT 14

S28994
antifungal protein 2 - white mustard (fragment)
C:Species: Sinapis alba (white mustard)
C:Date: 25-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 09-Jun-2000
C:Accession: S28994
R:Terras, F.R.G.; Torrekens, S.; van Leuven, F.; Osborn, R.W.; Vanderleyden, J.; Cammue, F.E.B. Lett. 316, 233-240, 1993
A;Title: A new family of basic cysteine-rich plant antifungal proteins from Brassicaceae
A;Reference number: S28989; MUID:93138130
A;Accession: S28994
A;Molecule type: protein
A;Residues: 1-26 <TER>
C;Superfamily: gamma-thionin
C;Keywords: phosphoprotein

Query Match 13.9%; Score 30; DB 2; Length 26;
Best Local Similarity 35.7%; Pred. No. 1.8e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 5 GTFSNTTSTTDICR 18
|||: | | | | |
Db 9 GTWSGVCNNACR 22

RESULT 15

C61233
conceptus protein 5 - cat (fragment)
C:Species: Felis silvestris catus (domestic cat)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 07-May-1999
C:Accession: C61233
R:Thatcher, M.J.D.; Shille, V.M.; Fliss, M.F.; Bazer, F.W.; Sisum, W.; Randal, S.

Biol. Reprod. 44, 108-120, 1991
A;Title: Characterization of feline conceptus proteins during pregnancy.
A;Reference number: A61233; MUID:91198359
A;Accession: C61233
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-29 <THA>
C;Superfamily: insulin-like growth factor binding protein 1; thyroglobulin type I rep

Query Match 13.9%; Score 30; DB 2; Length 29;
Best Local Similarity 20.0%; Pred. No. 2e+03;
Matches 7; Conservative 5; Mismatches 11; Indels 12; Gaps 1;

OY 2 CAPGTFSNTTSTDCRPHQICNVVAIPGNASMDA 36
|||: | | | | |
Db 7 CAP-----CSPEKALCPVPDSCQTQSA 29

Search completed: July 13, 2001, 17:14:48
Job time: 148 sec

Result No.	Score	Query %			DB	ID	Description
		Match	Length	DB			
1	38	17.6	33	1	TXHL_SFLHU	P56676 selenocosmi	
2	37.5	17.4	37	1	TXJG_HADVE	P82228 hadronyche	
3	37	17.1	36	1	TXJA_HADVE	P82227 hadronyche	
4	34	15.7	36	1	TXJB_HADVE	P82226 hadronyche	
5	33.5	15.5	39	1	PA2_HELHO	P04362 heloderma h	
6	32.5	15.0	35	1	TXAG_AGEOP	P31328 agelena opu	
7	31	14.4	38	1	PPOX_BOVIN	P56602 bos taurus	
8	31	14.4	30	1	CYOL_VLOOD	P82230 viola odora	
9	30	13.9	31	1	CYLA_PSYLO	P56872 psychotria	
10	30	13.9	38	1	BD01_BOVIN	P46159 bos taurus	
11	29	13.4	15	1	IDIH_PSESP	P80701 pseudomonas	
12	29	13.4	24	1	LEC_CROJU	P16352 crotalaria	
13	29	13.4	34	1	RR2_OCHNE	Q40606 ochrosphaer	
14	28	13.0	18	1	CXAL_CONER	P50982 conus ermin	
15	28	13.0	29	1	GLUC_PLAFE	P23062 platichthys	
16	28	13.0	30	1	IT1I_LAGLE	P26771 lagenaria l	
17	28	13.0	30	1	ITR1_CITLA	P1969 citrullus l	
18	28	13.0	36	1	GLU1_ORENI	P81026 oreochromis	
19	27.5	12.7	31	1	CXD6_CONNI	P56710 conus nigro	
20	27	12.5	22	1	CXM1_CONGE	P01523 conus geogr	
21	27	12.5	23	1	AFP3_BRANA	P30226 brassica na	
22	27	12.5	24	1	RS13_THETH	P80377 thermus aqu	
23	27	12.5	25	1	ANDT_ANDAU	P56684 androctonus	
24	27	12.5	27	1	AFP1_BRARA	P30227 brassica ra	
25	27	12.5	28	1	ETX2_BACCE	P80568 bacillus ce	
26	27	12.5	28	1	PA2C_PSEPO	P20260 pseudechis	
27	27	12.5	30	1	VPU_HVISC	P05948 human immu	
28	27	12.5	32	1	LPV1_ECOLI	P03061 escherichia	
29	26.5	12.3	32	1	IAPP_SHEEP	Q28605 ovis aries	
30	26	12.0	28	1	NLTX_WHEAT	P39085 triticum ae	
31	26	12.0	30	1	GLUW_ANGAN	P41521 anguilla an	
32	26	12.0	30	1	ITR1_WOMCH	P10294 momordica c	
33	26	12.0	34	1	DEF2_RABIT	P07468 cryctolagus	

This SWISS-PROT entry is copyright. It is produced through a collaboration

ALIGNMENTS

CC RECEPTOR: BLOCKS NEUROMUSCULAR TRANSMISSION.
CC -----

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AFI57504; AAF25774.1; -
 DR PDB; 1OK6; 20-AUG-99.
 KW Venom; Neurotoxin; Postsynaptic neurotoxin; 3D-structure.
 FT DISULFID 2 17
 FT DISULFID 9 22
 FT DISULFID 16 29
 SQ SEQUENCE 33 AA; 3756 MW; 1CCE219FD6D31F11 CRC64;

Query Match 17.6%; Score 38; DB 1; Length 33;
 Best Local Similarity 30.0%; Pred. No. 65;
 Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 5 GTFSTNTSSDTCRPHQICN 24
 Db 4 GVFDACTPGRNECCPNRVCS 23

RESULT 2
 TXJC_HADVE STANDARD; PRT; 37 AA.
 ID TXJC_HADVE
 AC P82228;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE J-ATRACOTOXIN-HV1C (J-ACTX-HV1C).
 OS Hadronyche versuta (Blue mountains funnel-web spider) (Atrax
 OS versutus).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
 OC Mygalomorphae; Hexathelidae; Hadronyche.
 OX NCBI_TaxID=6904;
 RN [1]
 RP SEQUENCE, FUNCTION, X-RAY CRYSTALLOGRAPHY, AND STRUCTURE BY NMR.

RC TISSUE-Venom gland;
 RX MEDLINE-20343014; PubMed=10881200;
 RA Wang X.-H., Connor M., Smith R., Maciejewski M.W., Howden M.E.H.,
 RA Nicholson G.M., Christie M.J., King G.F.;
 RT "Discovery and characterization of a family of insecticidal
 RT neurotoxins with a rare vicinal disulfide bridge."
 RL Nat. Struct. Biol. 7:505-513(2000).
 CC -1- FUNCTION: INSECTICIDAL NEUROTOXIN.

DR PDB; 1DL0; PRELIMINARY.
 KW Venom; Toxin; Neurotoxin; 3D-structure.
 FT DISULFID 3 17
 FT DISULFID 10 22
 FT DISULFID 13 14
 FT DISULFID 16 32
 SQ SEQUENCE 37 AA; 3768 MW; E4DDF046CC750FFC CRC64;

Query Match 17.4%; Score 37.5; DB 1; Length 37;
 Best Local Similarity 42.1%; Pred. No. 84;
 Matches 8; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 1 PCAPGTFSTNTSS-TDTCR 18
 Db 15 PCCPGTSCKAESNGSVYCR 33

RESULT 3
 TXJA_HADVE STANDARD; PRT; 36 AA.
 ID TXJA_HADVE
 AC P82227;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE J-ATRACOTOXIN-HV1A (J-ACTX-HV1A).
 OS Hadronyche versuta (Blue mountains funnel-web spider) (Atrax
 OS versutus).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
 OC Mygalomorphae; Hexathelidae; Hadronyche.
 OX NCBI_TaxID=6904;
 RN [1]
 RP SEQUENCE.

RC TISSUE-Venom gland;
 RX MEDLINE-20343014; PubMed=10881200;
 RA Wang X.-H., Connor M., Smith R., Maciejewski M.W., Howden M.E.H.,
 RA Nicholson G.M., Christie M.J., King G.F.;
 RT "Discovery and characterization of a family of insecticidal
 RT neurotoxins with a rare vicinal disulfide bridge."
 RL Nat. Struct. Biol. 7:505-513(2000).
 CC -1- FUNCTION: INSECTICIDAL NEUROTOXIN.

KW Venom; Toxin; Neurotoxin.
 FT DISULFID 3 17
 FT DISULFID 10 22
 FT DISULFID 13 14
 FT DISULFID 16 33
 SQ SEQUENCE 36 AA; 3685 MW; D1598B2560BFE997 CRC64;

Query Match 17.1%; Score 37; DB 1; Length 36;
 Best Local Similarity 40.0%; Pred. No. 96;
 Matches 8; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

QY 1 PCAPGTFSTNTSSDTCR 18
 Db 15 PCCPGTSQGPESNGVYCR 34

RESULT 4
 TXJB_HADVE STANDARD; PRT; 36 AA.
 ID TXJB_HADVE
 AC P82226;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE J-ATRACOTOXIN-HV1B (J-ACTX-HV1B).
 OS Hadronyche versuta (Blue mountains funnel-web spider) (Atrax
 OS versutus).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
 OC Mygalomorphae; Hexathelidae; Hadronyche.
 OX NCBI_TaxID=6904;
 RN [1]
 RP SEQUENCE.

RC TISSUE-Venom gland;
 RX MEDLINE-20343014; PubMed=10881200;
 RA Wang X.-H., Connor M., Smith R., Maciejewski M.W., Howden M.E.H.,
 RA Nicholson G.M., Christie M.J., King G.F.;
 RT "Discovery and characterization of a family of insecticidal
 RT neurotoxins with a rare vicinal disulfide bridge."
 RL Nat. Struct. Biol. 7:505-513(2000).
 CC -1- FUNCTION: INSECTICIDAL NEUROTOXIN.

KW Venom; Toxin; Neurotoxin.
 FT DISULFID 3 17
 FT DISULFID 10 22
 FT DISULFID 13 14
 FT DISULFID 16 33
 SQ SEQUENCE 36 AA; 3651 MW; D23A442560B89997 CRC64;

Query Match 15.7%; Score 34; DB 1; Length 36;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PCAPGT 6
 Db 15 PCCPGT 20

```

RESULT 5
PA2_HELHO STANDARD; PRT; 39 AA.
AC P04362;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PHOSPHOLIPASE A2 (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE)
DE (FRAGMENT).
OS Heloderma horridum horridum (Mexican beaded lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC Heloderma.
OX NCBI_TaxID=8552;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=86243292; PubMed=3087412;
RA Sosa B.P., Alagon A.C., Martin B.M., Possani L.D.;
RT "Biochemical characterization of the phospholipase A2 purified from
RT the venom of the Mexican beaded lizard (Heloderma horridum horridum
RT Wiegmann).";
RL Biochemistry 25:2927-2933(1986).
CC -1- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERYL-
CC -1- PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
CC PIR: A05323; A05323.
DR HSSP; P00630; IPOC.
DR InterPro: IPR001211.
DR PROSITE: PS00119; PA2_ASP; PARTIAL.
DR PROSITE: PS00118; PA2_HIS; 1.
KW Hydrolase; Lipid degradation; Calcium; Venom.
FT ACT_SITE 36 36 BY SIMILARITY.
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4179 MW; 8F9BC6B5DFB603E CRC64;

Query Match 15.5%; Score 33.5; DB 1; Length 39;
Best Local Similarity 31.0%; Pred. No. 3e+02;
Matches 9; Conservative 3; Mismatches 10; Indels 7; Gaps 2;

Qy 2 CAPGTFSN-----TTSSTDI-CRPHQIC 23
| | | | | | | | | | | | | | | | | | | | |
Db 11 CGAGNAASDYSQLGTEKDTMCCRHDHC 39

RESULT 6
TXAG_AGEOP STANDARD; PRT; 35 AA.
AC P31328;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE AGELENIN.
OS Agelena opulenta (Spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Agelenidae; Agelena.
OX NCBI_TaxID=29934;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA Hagiwara K., Sakai T., Miwa A., Kawai N., Nakajima T.;
RT "Complete amino acid sequence of a new type of neurotoxin from the
RT venom of the spider, Agelena opulenta.";
RL Biomed. Res. 11:181-186(1990).
RN [2]
RP DISULFIDE BONDS, AND AMIDATION.
RC TISSUE=Venom;
RA Hagiwara K., Inui T., Nakajima K., Kimura T., Kitada C., Fujino M.,
RA Sakakibara S., Nakajima T.;
RT "Agelenin, a spider neurotoxin: determination of the C-terminus as

```

```

RT amide form, and investigation of the disulfide bond arrangement.";
RL Biomed. Res. 12:357-363(1991).
RN [3]
RP SYNTHESIS, DISULFIDE BONDS, AND AMIDATION.
RC TISSUE=Venom;
RX MEDLINE=93043890; PubMed=1421801;
RA Inui T., Hagiwara K., Nakajima K., Kimura T., Nakajima T.,
RA Sakakibara S.;
RT "Synthesis and disulfide structure determination of agelenin:
RT identification of the carboxy-terminus as an amide form.";
RL Pept. Res. 5:140-144(1992).
CC -1- FUNCTION: SUPPRESSES THE EXCITATORY POSTSYNAPTIC POTENTIALS
CC POSSIBLY BY BLOCKING THE PRESYNAPTIC CALCIUM CHANNEL.
CC PIR: A60959; A60959.
DR Venom; Neurotoxin; Calcium channel inhibitor; Amidation.
KW DISULFID 3 19
FT DISULFID 10 24
FT DISULFID 18 34
FT MOD_RES 35 35
SQ SEQUENCE 35 AA; 3825 MW; CBE6462825350D90 CRC64;

Query Match 15.0%; Score 32.5; DB 1; Length 35;
Best Local Similarity 42.9%; Pred. No. 3.7e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 17 CRPH-QICNVVAIP 29
| | | | | | | | | | | | | | | | | | | | |
Db 3 CLPHNRCNALSGP 16

RESULT 7
PPOX_BOVIN STANDARD; PRT; 28 AA.
AC P56602;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO) (FRAGMENTS).
GN PPOX.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95331315; PubMed=7607249;
RA Taketani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,
RA Nishimura K., Inokuchi H.;
RT "Induction of terminal enzymes for heme biosynthesis during
RT differentiation of mouse erythroleukemia cells.";
RL Eur. J. Biochem. 230:760-765(1995).
CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC IX TO FORM PROTOPORPHYRIN IX.
CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) -> PROTOPORPHYRIN-
CC IX + H(2)O(2).
CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW Flavoprotein; FAD; Mitochondrion.
FT NON_TER 1 1
FT NON_CONS 13 14
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 2777 MW; FEDFC3F09CB6A345 CRC64;

Query Match 14.4%; Score 31; DB 1; Length 28;

```

```

Best Local Similarity 42.9%; Pred. No. 4.6e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 25 VVAIPGNASMDAVC 38
    | | | | | | | |
Db 4 VVVLGGISGDSLC 17

RESULT 8
CYOL_VIOOD STANDARD; PRT; 30 AA.
AC P82230;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE CYCLOVIOACIN OL.
OS Viola odorata (Sweet violet).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Malpighiales; Violaceae; Viola.
OX NCBI_TaxID=97441;
RN [1]
RP SEQUENCE.
RX MEDLINE=20069951; PubMed=10600388;
RA Craik J.D., Daly N.D., Bond T., Waine C.;
RT "Plant cyclotides - a unique family of cyclic and knotted proteins
    that defines the cyclic cysteine knot structural motif.";
RL J. Mol. Biol. 294:1327-1336(1999).
CC -1- PTM: THIS IS A CYCLIC PEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE CYCLOTIDE FAMILY.
CC -1- CAUTION: THIS PEPTIDE BEING CYCLIC, ITS SEQUENCE WAS ARBITRARILY
    CHOSEN TO START AT THE POSITION SHOWN BELOW, AS THE DNA SEQUENCE
    FROM WHICH IT IS DERIVED IS NOT YET KNOWN, IT IS NOT POSSIBLE TO
    ASSIGN THE CORRECT N- AND C-TERMINI.
FT DISULFID 2 17
FT DISULFID 7 22
FT DISULFID 15 28
SQ SEQUENCE 30 AA; 3141 MW; A3D8CA231098E7E7 CRC64;

Query Match 14.4%; Score 31; DB 1; Length 30;
Best Local Similarity 37.5%; Pred. No. 5e+02;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 23 CNVVAIPGNASMDAVC 38
    | | | | | | | |
Db 7 CTVTALLGSCSNRVC 22

RESULT 9
CYLA_PSYLO STANDARD; PRT; 31 AA.
AC P56872; P82254;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE CYCLOPSYCHOTRIDE A (CPT).
OS Psychotria longipes.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Gentianales; Rubiaceae; Psychotria.
OX NCBI_TaxID=41680;
RN [1]
RP SEQUENCE.
RX MEDLINE=95230294; PubMed=7714530;
RA Witherup K.M., Bogusky M.J., Anderson P.S., Ramjit H., Ransom R.W.,
RA Wood T., Sardana M.;
RT "Cyclopsychotride A, a biologically active, 31-residue cyclic peptide
    isolated from Psychotria longipes.";
RL J. Nat. Prod. 57:1619-1625(1994).
RN [2]
RP SYNTHESIS, AND ANTIBACTERIAL ACTIVITY.
RX MEDLINE=99362685; PubMed=10430870;

```

```

RA Tam J.P., Lu Y.-A., Yang J.-L., Chiu K.-W.;
RT "An unusual structural motif of antimicrobial peptides containing
    end-to-end macrocycle and cystine-knot disulfides.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:8913-8918(1999).
CC -1- FUNCTION: PROBABLY PARTICIPATES IN A PLANT DEFENSE MECHANISM. HAS
    ANTIBIOTIC ACTIVITY. INHIBITS NEUTROSENSIN BINDING. ACTIVE AGAINST
    BOTH GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
CC -1- PTM: THIS IS A CYCLIC PEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE CYCLOTIDE FAMILY.
CC -1- CAUTION: THIS PEPTIDE BEING CYCLIC, ITS SEQUENCE WAS ARBITRARILY
    CHOSEN TO START AT THE POSITION SHOWN BELOW, AS THE DNA SEQUENCE
    FROM WHICH IT IS DERIVED IS NOT YET KNOWN, IT IS NOT POSSIBLE TO
    ASSIGN THE CORRECT N- AND C-TERMINI.
*FW Antibiotic.
KW DISULFID 2 17 BY SIMILARITY.
FT DISULFID 7 22 BY SIMILARITY.
FT DISULFID 15 29 BY SIMILARITY.
SQ SEQUENCE 31 AA; 3255 MW; ACC0BBB232ED0CD0 CRC64;

Query Match 13.9%; Score 30; DB 1; Length 31;
Best Local Similarity 37.5%; Pred. No. 7e+02;
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 23 CNVVAIPGNASMDAVC 38
    | | | | | | | |
Db 7 CTVTALLGSCSKVC 22

RESULT 10
BD01_BOVIN STANDARD; PRT; 38 AA.
AC P46159;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BETA-DEFENSIN 1 (BNDB-1) (BNDB-1).
GN DEFB1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX STRAIN=HEREFORD; TISSUE=Neutrophils;
RX MEDLINE=93203264; PubMed=8454635;
RA Selsted M.E., Tang Y.-Q., Morris W.L., McGuire P.A., Novotny M.J.,
RA Smith W., Henschen A.H., Cullor J.S.;
RT "Purification, primary structures, and antibacterial activities of
    beta-defensins, a new family of antimicrobial peptides from bovine
    neutrophils.";
RL J. Biol. Chem. 268:6641-6648(1993).
CC -1- FUNCTION: HAS BACTERICIDAL ACTIVITY. ACTIVE AGAINST E. COLI ML35
    BUT NOT AGAINST S. AUREUS 502A.
CC -1- TISSUE SPECIFICITY: NEUTROPHILIC GRANULES.
CC -1- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
DR HSP; P46170; I8NB.
DR InterPro; IPR001855; -.
DR Pfam; PF00711; Defensin_beta; 1.
KW Antibiotic.
FT DISULFID 5 34 BY SIMILARITY.
FT DISULFID 12 27 BY SIMILARITY.
FT DISULFID 17 35 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4278 MW; 48B872D1025E1A68 CRC64;

Query Match 13.9%; Score 30; DB 1; Length 38;
Best Local Similarity 21.9%; Pred. No. 8.5e+02;
Matches 7; Conservative 8; Mismatches 11; Indels 6; Gaps 1;

QY 7 FSNNTSSTDICPHQICNVVAIPGNASMDAVC 38
    | | | | | | | |

```

Db 2 FASCHTNGIGICLPNR-----CPGHMIGIC 27

RESULT 11
DIDH_PSESP STANDARD; PRT; 15 AA.
AC P80701;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 3-ALPHA-HYDROXYSTEROID DEHYDROGENASE (EC 1.1.1.50) (3-ALPHA-HSD)
DE (HYDROXYPROSTAGLANDIN DEHYDROGENASE) (HSD29) (FRAGMENT).
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE.
RX MEDLINE=97100200; PubMed=8944761;
RA Oppermann U.C.T., Maser E.;
RT "Characterization of a 3 alpha-hydroxysteroid dehydrogenase/carbonyl
RL Eur. J. Biochem. 241:744-749(1996).
CC -1- FUNCTION: ALONG WITH THE 3 ALPHA-HYDROXYSTEROID DEHYDROGENASE AND
CC 3-OXO-REDUCTASE ACTIVITIES TOWARDS A VARIETY OF CIS OR TRANS FUSED
CC A/B RING STEROIDS, IT ALSO REDUCES SEVERAL XENOBIOTIC CARBONYL
CC COMPOUNDS, INCLUDING A METRAPONE-BASED CLASS OF INSECTICIDES, TO
CC THE RESPECTIVE ALCOHOL METABOLITES.
CC -1- CATALYTIC ACTIVITY: ANDROSTERONE + NAD(P)(+) =
CC 5-ALPHA-ANDROSTANE-3,17-DIONE + NAD(P)H.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR InterPro; IPR002198; -;
DR PROSITE; PS00061; ADH_SHORT; PARTIAL.
KW Oxidoreductase; NAD.
FT DOMAIN 6 >15 INVOLVED IN COFACTOR BINDING
FT (BY SIMILARITY).
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1315 MW; 9506860D070A7790 CRC64;
Query Match 13.4%; Score 29; DB 1; Length 15;
Best Local Similarity 66.78; Pred. No. 4.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 25 VVAIPGNAS 33
I: I I I I I
Db 2 VIAITGSAS 10

RESULT 12
LEC_CROJU STANDARD; PRT; 24 AA.
AC P16352;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE LECTIN (FRAGMENT).
OS Crotalaria juncea (Sunn hemp).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Crotalaria.
OX NCBI_TaxID=3829;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RA Foriers A., de Neve R., Strosberg A.D.;
RT "Lectin sequences as a tool for chemotaxonomical classification.";,
RL physiol. Veg. 17:597-606(1979).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
DR PR; S08293; S08293.
DR HSP; P04122; ILGB.

DR InterPro; IPR000985; -;
DR InterPro; IPR001220; -;
DR Pfam; PF00139; lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; PARTIAL.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; PARTIAL.
KW Lectin; Glycoprotein.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2614 MW; 80704D8CD9F9BB6E CRC64;
Query Match 13.4%; Score 29; DB 1; Length 24;
Best Local Similarity 60.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 6 TFSNTSTSD 15
I: I I I I I
Db 5 SFSSTKSTSD 14

RESULT 13
RR2_OCHNE STANDARD; PRT; 34 AA.
AC Q40606;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S2 (FRAGMENT).
GN RPS2.
OS Ochrosphaera neapolitana.
OC Chloroplast.
OC Eukaryota; Haptophyceae; Isochrysidales; Ochrosphaera.
OX NCBI_TaxID=35137;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCMP 593;
RA Huss V.A.R., Tietze A.C., Julius C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99078; CAA67534.1; -;
DR InterPro; IPR001865; -;
DR Pfam; PF00318; Ribosomal_S2; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; PARTIAL.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; PARTIAL.
KW Ribosomal protein; Chloroplast.
FT NON_TER 1 1
SQ SEQUENCE 34 AA; 3640 MW; 2039BA0FB5710655 CRC64;
Query Match 13.4%; Score 29; DB 1; Length 34;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Qy 24 NVVAIPGNASMDAV 37
I: I I I I I
Db 1 NLVDIPANDDAI 14

RESULT 14
CXAL_CONER STANDARD; PRT; 18 AA.
AC P50982;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE ALPHA-CONOTOXIN EI.
 OS Conus ermineus (Atlantic fish-hunting cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=55423;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=96062516; PubMed=7578057;
 RA Martinez J.S., Olivera B.M., Gray W.R., Craig A.G., Groebe D.R.,
 RA Abramson S.N., McIntosh J.M.;
 RT "Alpha-conotoxin EI, a new nicotinic acetylcholine receptor
 RT antagonist with novel selectivity."
 RL Biochemistry 34:14519-14526(1995).
 CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 CC INHIBIT THEM.
 CC Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
 KW Venom; Hydroxylation.
 FT DISULFID 4 10
 FT DISULFID 5 18
 FT MOD.RES 3 3 HYDROXYLATION.
 FT MOD.RES 18 18 AMIDATION
 SQ SEQUENCE 18 AA; 2082 MW; 60A61A6C427A6B5E CRC64;

Query Match 13.0%; Score 28; DB 1; Length 18;
 Best Local Similarity 45.5%; Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 15 DICRPHQICNV 25
 | | | | |
 Db 2 DPCCYHPTCNM 12

RESULT 15
 GLUC_PLAFE STANDARD; PRT; 29 AA.
 AC P23062;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE GLUCAGON.
 OS Platicthys flesus (European flounder), and
 OS Thunnus obesus (Bigeye tuna).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectoidei; Pleuronectidae; Platicthys.
 OX NCBI_TaxID=8260, 8241;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=P.flesus;
 RX MEDLINE=87219793; PubMed=3556313;
 RA Conlon J.M., Davis M.S., Thim L.;
 RT "Primary structure of insulin and glucagon from the flounder
 RT (Platicthys flesus).";
 RL Gen. Comp. Endocrinol. 83:227-232(1991).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=T.obesus; TISSUE=Pancreas;
 RX MEDLINE=92009094; PubMed=1916209;
 RA Navarro I., Gutierrez J., Caixach J., Rivera J., Planas J.;
 RT "Isolation and primary structure of glucagon from the endocrine
 RT pancreas of Thunnus obesus."
 RL Gen. Comp. Endocrinol. 83:227-232(1991).
 CC -|- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -|- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -|- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR; S09348; GCFLE.
 DR PIR; A61135; A61135.

DR HSSP; P01274; IGCN.
 DR InterPro; IPR000532;
 DR Pfam; PF00123; hormone2; 1.
 DR PRINTS; PR00275; GLUCAGON.
 DR PROSITE; PS00260; GLUCAGON; 1.
 KW Glucagon family; Hormone.
 SQ SEQUENCE 29 AA; 3508 MW; 77D5943208662E52 CRC64;

Query Match 13.0%; Score 28; DB 1; Length 29;
 Best Local Similarity 75.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GTFSNTTS 12
 | | | | |
 Db 4 GTFSNDYS 11

Search completed: July 13, 2001, 17:16:26
 Job time: 166 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: July 13, 2001, 17:13:25 ; Search time 20.87 Seconds
(without alignments)
247.240 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_201
Perfect score: 216
Sequence: 1 PCAPGTFSTTSSTDICRPHQICNVVAIPGNASMDAVCT 39

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 18182

```

Minimum DB seq length: 0
Maximum DB seq length: 39

```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
SPTREMBL_16.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_unclassified.*
13: sp_vertebrate.*
14: sp_virus.*
```

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	121	56.0	30	4	Q9UHL1	Q9uihl homo sapien
2	38.5	17.8	37	4	Q9UBD4	Q9ubd4 homo sapien
3	36	16.7	35	14	Q70293	Q70293 human immun
4	36	16.7	35	14	Q70294	Q70294 human immun
5	36	16.7	35	14	Q70295	Q70295 human immun
6	36	16.7	35	14	Q70351	Q70351 human immun
7	35	16.2	26	9	Q9ZXH9	Q9zxh9 bacteriopho
8	34.5	16.0	36	6	P79330	P79330 bos taurus
9	34	15.7	33	14	Q9IEX1	Q9iex1 cotton leaf
10	33	15.3	15	11	Q9QUY5	Q9quy5 rattus sp.
11	33	15.3	39	2	Q9PBZ7	Q9pbz7 xylella fas
12	32	14.8	33	4	Q9UD12	Q9ud12 homo sapien
13	32	14.8	34	14	Q9QL92	Q9ql92 human adeno
14	31.5	14.6	29	6	Q9MZ98	Q9mz98 pongo pygma
15	31.5	14.6	33	3	Q9HFZ6	Q9hfh2 cryptococcu
16	31.5	14.6	33	3	Q9HDP1	Q9hdp1 filobasidie
17	31.5	14.6	33	3	Q9HDP0	Q9hdp0 cryptococcu
18	31	14.4	29	4	Q9TRD4	Q9trd4 cryptolagus
19	31	14.4	29	5	Q9N3L0	Q9n3l0 caenorhabdi

20	31	14.4	30	14	Q77939	Q77939 human immun
21	31	14.4	31	4	O76030	O76030 homo sapien
22	31	14.4	33	3	Q9HF27	Q9HF27 cryptococcu
23	31	14.4	38	4	Q9NOY1	Q9NOY1 homo sapien
24	31	14.4	38	8	Q9T2N6	Q9T2N6 spinacia ol
25	30.5	14.1	30	6	Q9MZX0	Q9MZX0 colobus pol
26	30	14.1	32	2	O05602	O05602 pseudomonas
27	30	13.9	20	14	Q78505	Q78505 human immun
28	30	13.9	26	11	Q35630	Q35630 mus musculus
29	30	13.9	27	14	Q91TU5	Q91TU5 hepatitis c
30	30	13.9	36	10	Q99214	Q99214 aegilops sq
31	30	13.9	37	2	Q9KYN6	Q9KYN6 streptomyce
32	30	13.9	39	4	Q16368	Q16368 homo sapien
33	30	13.9	39	4	O13254	O13254 homo sapien
34	30	13.9	39	14	Q10486	Q10486 human immun
35	30	13.9	39	14	Q9QQT2	Q9QQT2 tanapox vir
36	29.5	13.7	22	3	Q9UR51	Q9UR51 filobasidie
37	29.5	13.7	26	2	O46711	O46711 plasmid r10
38	29.5	13.7	33	11	Q9WU11	Q9WU11 mus musculus
39	29.5	13.7	34	14	Q73447	Q73447 human papil
40	29.5	13.7	35	14	Q9QFA0	Q9QFA0 human immun
41	29.5	13.7	36	14	O89539	O89539 human papil
42	29.5	13.7	37	14	O41433	O41433 human immun
43	29.5	13.7	39	14	O10485	O10485 human immun
44	29	13.4	17	14	Q78381	Q78381 human immun
45	29	13.4	18	14	Q98EX3	Q98EX3 human immun

ALIGNMENTS

```

RESULT 1
Q9UIHI ID Q9UIHI PRELIMINARY; PRT; 30 AA.
AC Q9UIHI;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 (FRAGMENT).
GN TNFR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Konata T., Tsuchiya N., Matsushita M., Tokunaga K.;
RT "New polymorphism within the extracellular region of TNFR2.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030950; BAA89053.1; -.
KW Receptor.
FT NON_TER 1
FT NON_TER 30
SQ SEQUENCE 30 AA; 3183 MW; 942C00239B909DF5 CRC64;

Query Match 56.0%; Score 121; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 PCAPGTFNSTTSSDIDCRPHQ 21
Db 10 PCAPGTFNSTTSSDIDCRPHQ 30

RESULT 2
Q9UDB4 ID Q9UDB4 PRELIMINARY; PRT; 37 AA.
AC Q9UDB4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CD8 BETA CHAIN ISOFORM S BETA5 (FRAGMENT).

```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93170376; PubMed=8436166;
RA DiSanto J.P., Smith D., de Bruin D., Lacy E., Flomenberg N.;
RT "Transcriptional diversity at the duplicated human CD8 beta loci.";
RL Eur. J. Immunol. 23:320-326(1993).
SQ SEQUENCE 37 AA; 3940 MW; 7A4AE1C901E15E99 CRC64;

Query Match 17.8%; Score 38.5; DB 4; Length 37;
Best Local Similarity 34.8%; Pred. No. 1.7e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 1 PCAPGTFSTNTSSTDICRPHIC 23
DB 13 PLSPNACMDTTA---ILQPHRSK 32

RESULT 3
QY Q70293 PRELIMINARY; PRT; 35 AA.
AC Q70293;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
RA Weber J.N., Cheingsong-Popov R.;
RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Callow D.,
RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
RA Weber J.N.;
RT "Identification of human immunodeficiency virus type 1 subtypes and
their distribution in the Commonwealth of Independent States (Former
Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
analysis.";
RL J. Infect. Dis. 168:292-297(1993).
DR EMBL; U10706; AAA19262.1; -.
DR InterPro; IPR000777; -.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 14; Length 35;
Best Local Similarity 47.6%; Pred. No. 3.7e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 2 CAPGTFSTNTSST--DICRPH 20
DB 14 CAPCAQFYTTGTEIGDIRRAH 34

RESULT 5
QY Q70295 PRELIMINARY; PRT; 35 AA.
AC Q70295;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
RA Weber J.N., Cheingsong-Popov R.;
RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Callow D.,
RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
RA Weber J.N.;
RT "Identification of human immunodeficiency virus type 1 subtypes and
their distribution in the Commonwealth of Independent States (Former
Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
analysis.";
RL J. Infect. Dis. 168:292-297(1993).
DR EMBL; U10705; AAA19262.1; -.
DR InterPro; IPR000777; -.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 14; Length 35;
Best Local Similarity 47.6%; Pred. No. 3.7e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 2 CAPGTFSTNTSST--DICRPH 20
DB 14 CAPCAQFYTTGTEIGDIRRAH 34

RESULT 4
QY Q70294 PRELIMINARY; PRT; 35 AA.
AC Q70294;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
RA Weber J.N., Cheingsong-Popov R.;
RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Callow D.,
RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
RA Weber J.N.;
RT "Identification of human immunodeficiency virus type 1 subtypes and
their distribution in the Commonwealth of Independent States (Former
Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
analysis.";
RL J. Infect. Dis. 168:292-297(1993).
DR EMBL; U10706; AAA19263.1; -.
DR InterPro; IPR000777; -.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;
```

RT "Identification of human immunodeficiency virus type 1 subtypes and
 RT their distribution in the Commonwealth of Independent States (Former
 RT Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
 RT analysis.";
 RL J. Infect. Dis. 168:292-297(1993).
 DR EMBL: U10707; AAA19264.1; -;
 DR InterPro: IPR000777; -;
 DR Pfam: PF00516; GP120; 1.
 KW Envelope protein.
 FT NON_TER 1
 FT NON_TER 35
 FT NON_TER 35
 SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 14; Length 35;
 Best Local Similarity 47.6%; Pred. No. 3.7e+02;
 Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

Qy 2 CAPGTFSTST--DICRPH 20
 |||| || || ||
 Db 14 CAPGQAFYTTGTEIGDIRRAH 34

RESULT 6

Q70351 PRELIMINARY; PRT; 35 AA.

AC Q70351;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94338597; PubMed=8060542;
 RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
 RA Weber J.N., Cheingsong-Popov R.;
 RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
 RT of env V3 sequences and their correlation with epidemiologic data.";
 RL AIDS 8:619-624(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93329178; PubMed=8335967;
 RA Cheingsong-Popov R., Bobkov A.F., Garaev M., Kaleebu P., Callow D.,
 RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
 RA Weber J.N.;
 RT "Identification of human immunodeficiency virus type 1 subtypes and
 RT their distribution in the Commonwealth of Independent States (Former
 RT Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
 RT analysis.";
 RL J. Infect. Dis. 168:292-297(1993).
 DR EMBL: U10764; AAA19329.1; -;
 DR InterPro: IPR000777; -;
 DR Pfam: PF00516; GP120; 1.
 KW Envelope protein.
 FT NON_TER 1
 FT NON_TER 35
 FT NON_TER 35
 SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 14; Length 35;
 Best Local Similarity 47.6%; Pred. No. 3.7e+02;
 Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

Qy 2 CAPGTFSTST--DICRPH 20
 |||| || || ||
 Db 14 CAPGQAFYTTGTEIGDIRRAH 34

RESULT 7

Q92XH9

ID Q92XH9 PRELIMINARY; PRT; 26 AA.
 AC Q92XH9;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE PUTATIVE VERTEX HEAD SUBUNIT (FRAGMENT).
 GN G24.
 OS Bacteriophage ARI.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
 OT T4-like phages.
 OX NCBI_TaxID=66711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ARI;
 RA Syu W.J., Ding H.C., Seah J.N., Wu K.M., Yu S.L., Tam M.F.,
 RA Chang Y.C.;
 RT "Characterization of ARI coliphage specific to enterohemorrhagic
 RT Escherichia coli O157:H7.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF022930; AAD01756.1; -;
 FT NON_TER 26
 FT NON_TER 26
 SQ SEQUENCE 26 AA; 2830 MW; 5875E0CFBB665934 CRC64;

Query Match 16.2%; Score 35; DB 9; Length 26;
 Best Local Similarity 50.0%; Pred. No. 3.8e+02;
 Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 8 SNTSTSDICRPHQICNVVA 27
 | ||| | || | || |
 Db 11 STTNSNSIGRP----NLVA 26

RESULT 8

P79330 PRELIMINARY; PRT; 36 AA.

ID P79330;
 AC P79330;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE AGGREGAN EPIDERMAL GROWTH FACTOR-LIKE DOMAIN-1 (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CARTILAGE PROTEOGLYCAN;
 RX MEDLINE=97079270; PubMed=8921002;
 RA Fulop C., Cs-Szabo G., Glant T.T.;
 RT "Species-specific alternative splicing of the epidermal growth factor-
 RT like domain 1 of cartilage aggrecan.";
 RL Biochem. J. 319:935-940(1996).
 DR EMBL: L29486; AAB48067.1; -;
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 36 AA; 3614 MW; 9F6C32F15D3BC9DB CRC64;

Query Match 16.0%; Score 34.5; DB 6; Length 36;
 Best Local Similarity 34.8%; Pred. No. 6.1e+02;
 Matches 8; Conservative 1; Mismatches 7; Indels 7; Gaps 1;

Qy 1 PCAPGTFSTSTSDICRPHQIC 23
 ||: || | || |
 Db 9 PCSAGTCQETEG-----HVIC 24

RESULT 9

Q91EX1 PRELIMINARY; PRT; 33 AA.

ID Q91EX1

AC Q91EX1;

```

DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE AC4 PROTEIN (FRAGMENT).
GN AC4.
OS cotton leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=53010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PI2-IR;
RX MEDLINE=20318672; PubMed=10859391;
RA Sanz A.I., Fraile A., Garcia-Arenal F., Zhou X., Robinson D.J.,
RA Khalid S., Butt T., Harrison B.D.;
RT "Multiple infection, recombination and genome relationships among
RT begomovirus isolates found in cotton and other plants in Pakistan.";
RL J. Gen. Virol. 81:1839-1849(2000).
DR EMBL: AJ270854; CAB97069.1; -.
DR InterPro: IPR002489; -.
DR InterPro: IPR002511; -.
DR Pfam: PF01492; Gemini_C4; 1.
DR ProDom: PD002978; -. 1.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3478 MW; F3121B92E34ED31E CRC64;

Query Match 15.7%; Score 34; DB 14; Length 33;
Best Local Similarity 41.7%; Pred. No. 6.6e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTSSSTDICRPH 20
Db 15 NSNAGTVLRPH 26

RESULT 10
Q9QUY5 PRELIMINARY; PRT; 15 AA.
AC Q9QUY5
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE OLIGODENDROCYTE-SPECIFIC UDP-GALACTOSE: CERAMIDE GALACTOSYLTRANSFERASE
DE (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=96085162; PubMed=8521863;
RA Schulte S., Stoffel W.;
RT "UDP galactose:ceramide galactosyltransferase and glutamate/aspartate
RT transporter. Copurification, separation and characterization of the
RT two glycoproteins.";
RL Eur. J. Biochem. 233:947-953(1995).
SQ SEQUENCE 15 AA; 1657 MW; 84474749A06BFFCC CRC64;

Query Match 15.3%; Score 33; DB 11; Length 15;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PGTFSTNTS 12
Db 2 PGIFXSTTS 10

RESULT 11
Q9PBZ7 PRELIMINARY; PRT; 39 AA.
ID Q9PBZ7
AC Q9PBZ7
DT 01-OCT-2000 (TReMBLrel. 15, Created)

```

```

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN XF1988.
GN XF1988.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tettore A.H.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL: AE004018; AAF84790.1; -.
KW Hypothetical protein.
SQ SEQUENCE 39 AA; 4501 MW; E085D64BE286D612 CRC64;

Query Match 15.3%; Score 33; DB 2; Length 39;
Best Local Similarity 37.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 6 TFSNTSSTDICRPHQ 21
Db 24 SLSKVTVSADVVRHR 39

RESULT 12
Q9UD12 PRELIMINARY; PRT; 33 AA.
ID Q9UD12
AC Q9UD12
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE AMGX PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95322983; PubMed=7599636;
RA Lench N.J., Winter G.B.;
RT "Characterisation of molecular defects in X-linked amelogenesis
RT imperfecta (AIH1).";
RL Hum. Mutat. 5:251-259(1995).

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 16:58:40 ; Search time 19.99 Seconds
(without alignments)
118.276 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_201

Perfect score: 216

Sequence: 1 PCAPGTFSTSTDCRPHQICNVVAIPGNASMDAVCT 39

Scoring table: BLOSUMP2

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 206972

Minimum DB seq length: 0

Maximum DB seq length: 39

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	46	21.3	17	AAW95323	Costant and variab
2	46	21.3	38	AAV12433	Human 5' EST secre
3	42	19.4	36	AA373394	Human secreted pep
4	41	19.0	17	AAV51961	P. pastoris lysyl
5	41	19.0	17	AAV51973	P. pastoris lysyl
6	40	18.5	28	AAV64941	Human 5' Est relat
7	39	18.1	26	AAV24434	Adenovirus hexon p
8	37.5	17.4	34	AA444864	Human secreted pro
9	37	17.1	22	AA451542	YadA homologous pe
10	37	17.1	23	AAW73416	Human secreted pro
11	36	16.7	22	AAW70346	NF-AT transcriptio

12	36	16.7	22	21	AAV96558	Human NF-ATc1 anti
13	36	16.7	22	22	AA666490	Hepatitis B surfac
14	36	16.7	23	19	AAW65481	Antigenic site of
15	36	16.7	23	20	AAW97502	Gene #21 associate
16	36	16.7	27	22	AA80444	[(21Cys, 22-32pept
17	36	16.7	28	14	AA37510	Hepatitis B surfac
18	36	16.7	32	19	AAW65475	Hepatitis B surfac
19	36	16.7	32	19	AAW65476	Hepatitis B surfac
20	36	16.7	32	19	AAW65479	Hepatitis B surfac
21	36	16.7	32	19	AAW65480	Hepatitis B surfac
22	35.5	16.4	38	20	AAW02252	A F-box protein se
23	35.5	16.4	17	11	AAW06092	Immunoreactive pep
24	35.5	16.4	37	21	AAV91614	Human secreted pro
25	35	16.2	23	21	AAW27643	Human secreted pro
26	35	16.2	25	19	AAW44915	Spacer peptide for
27	35	16.2	27	17	AAW04374	Mu-conotoxin precu
28	35	16.2	31	6	AAW50156	Sequence of synthe
29	35	16.2	35	19	AAW44920	Spacer peptide for
30	35	16.2	36	18	AAW01811	Human papillomavir
31	34.5	16.0	30	15	AAW53570	Spider venom calcl
32	34	15.7	26	15	AAW47423	PDGF-activity-disp
33	34	15.7	30	20	AAW92254	BL172 polypeptide
34	34	15.7	36	14	AAW39319	Fla toxin. Atrax
35	34	15.7	36	19	AAW41884	Peptide used in ra
36	34	15.7	36	22	AAW61439	Human TANGO 275 EG
37	34	15.7	37	14	AAW39318	Fla toxin. Atrax
38	34	15.7	37	17	AAW88816	Respiratory syncyt
39	34	15.7	37	20	AAV13109	Human secreted pro
40	34	15.7	38	18	AAW12748	A-lineage conotoxi
41	34	15.7	38	20	AAW02251	A F-box protein se
42	33.5	15.5	16	19	AAW37752	Antigenic C-termin
43	33.5	15.5	20	21	AAW23019	Human APC protein
44	33.5	15.5	35	21	AAW89114	Core polypeptide f
45	33.5	15.5	35	21	AAW89115	Core polypeptide f

ALIGNMENTS

RESULT 1

AAW95323

ID AAW95323 standard; Protein; 17 AA.

XX AC AAW95323;

XX DT 15-MAR-1999 (first entry)

XX DE Costant and variable domain sequence of C. psittaci CPS92-106.

XX Chlamydia; cryptic phase; elementary body phase; replicating; probenidicid;

KW antiporphyrin acid; immune response; infection; diagnostic; assay; MOMP;

KW major outer membrane protein; autoimmune; inflammatory; porphyria;

KW Ebstein Barr virus; antioxidant.

XX OS Chlamydia psittaci.

XX PN WO9850074-A2.

XX PD 12-NOV-1998.

XX PF 06-MAY-1998; 98WO-US09237.

XX PR 18-FEB-1998; 98US-0025521.

XX PR 06-MAY-1997; 97US-0045689.

XX PR 06-MAY-1997; 97US-0045739.

XX PR 06-MAY-1997; 97US-0045779.

XX PR 06-MAY-1997; 97US-0045780.

XX PR 06-MAY-1997; 97US-0045784.

XX PR 06-MAY-1997; 97US-0045787.

XX PR 14-AUG-1997; 97US-0911593.

XX PR 18-FEB-1998; 98US-0025174.

XX PR 18-FEB-1998; 98US-0025176.

PA (UYVA-) UNIV VANDERBILT.
XX Mitchell WM, Stratton CW;
XX WPI; 1999-059653/05.
XX Composition with two agents effective against different stages of
PT Chlamydial life cycle - comprises agent targetted against cryptic
PT phase, against elementary body phase, against replicating phase,
PT probenicid and antiporphyric
XX Claim 4; Fig 3; 138pp; English.
XX The invention relates to the diagnosis and management of infections by
CC Chlamydia species. The invention provides a composition that comprises
CC at least two agents, where each of the agents is effective against a
CC different phase of the chlamydial life cycle. The agents are selected
CC from: (a) agents targetted against cryptic phase of chlamydial life
CC cycle; (b) agents targetted against elementary body phase of chlamydial
CC life cycle; (c) agents targetted against replicating phase of chlamydial
CC life cycle; (d) probenicid, and (e) antiporphyric acid. The composition
CC is used to elicit a protective immune response to Chlamydia infection in
CC an animal or human and is applied until the animal or human tests
CC negative for Chlamydia infection. It is also used to treat biological
CC material infected with Chlamydia. Diagnostic kits for antibody assays
CC against recombinant major outer membrane protein (MOMP), and for DNA
CC amplification assays for chlamydial genes, are used to diagnose disease,
CC e.g. autoimmune disease, an inflammatory disease or a disease that
CC occurs in an immuno-compromised individual, associated with Chlamydia
CC infection. The kits are used to detect chlamydial elementary bodies in a
CC sample. They are also used to monitor and/or modify the course of therapy
CC in a patient. The treatment reduces the acellular load of infectious
CC Ebsstein Barr virus. The method is also used to treat porphyria, by
CC reducing the number of elementary bodies and applying a drug, e.g.
CC cimetidine, and antioxidants, to reduce the adverse effects associated
CC with porphyria. Sequences AAW95320 to AAW95323 represent constant and
CC variable domain sequences of various Chlamydia species.
XX Sequence 17 AA;
SQ

Query Match 21.3%; Score 46; DB 20; Length 17;
Best Local Similarity 64.3%; Pred. No. 9.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 CAPGTFSTNTSSSTD 15
Db 1 casgtasnttvaad 14
II III IIII : I
1 casgtasnttvaad 14

RESULT 2
AAY12433
ID AAY12433 standard; Protein; 38 AA.
XX
XX AAY12433;
XX
XX 17-JUN-1999 (first entry)
DT
DE Human 5' EST secreted protein SEQ ID NO:464.
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductively hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
OS
XX WO9906548-A2.
PN
XX 11-FEB-1999.
PD
XX

PF 31-JUL-1998; 98WO-IB01222.
XX
PR 01-AUG-1997; 97US-0905135.
XX
XX (GEST) GENSET.
PA
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX WPI; 1999-153778/13.
XX N-PSDB; AAX41266.
XX New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT kidney, lung, umbilical cord, placenta and colon tissue
XX Claim 27; Page 768-769; 824pp; English.
XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, chemotactic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX Sequence 38 AA;
SQ

Query Match 21.3%; Score 46; DB 20; Length 38;
Best Local Similarity 40.9%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 2 CAPGTFSTNTSSSTDICRPHQIC 23
Db 10 crpatldqatratpcrlsqqc 31
II III IIII : I
10 crpatldqatratpcrlsqqc 31

RESULT 3
AAB37394
ID AAB37394 standard; Peptide; 36 AA.
XX
XX AAB37394;
AC
XX 20-FEB-2001 (first entry)
DT
XX Human secreted peptide #32 encoded by CDNA #47.
DE
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
XX Homo sapiens.
OS
XX WO200058335-A1.
PN
XX 05-OCT-2000.
PD
XX 22-MAR-2000; 2000WO-US07534.
PF
XX 26-MAR-1999; 99US-0126598.
PR
XX 22-DEC-1999; 99US-0171504.
PR

CC peroxidases, lysyl oxidases, protein disulfide reductases, tyrosine
 CC oxidases or sulphydryl oxidases is used to formulate compositions
 CC containing active ingredients. (I) is useful for crosslinking protein
 CC layers surrounding active ingredients in food, animal feed and
 CC pharmaceutical products. The compositions can be formulated without
 CC using chemical crosslinking agents. AAY51963-Y51974 represent fragments
 CC of the Pichia pastoris lysyl oxidase protein which is used to
 CC illustrate the method of the invention.
 XX
 SQ Sequence 17 AA;

Query Match 19.0%; Score 41; DB 21; Length 17;
 Best Local Similarity 70.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNT 10
 ||||| ||
 Db 7 pcapgvvynt 16

RESULT 6
 AAY64941
 ID AAY64941 standard; Protein; 28 AA.

XX AC AAY64941;

XX DT 01-FEB-2000 (first entry)

XX DE Human 5' EST related polypeptide SEQ ID NO:1102.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification.

XX OS Homo sapiens.

XX PN WO9953051-A2.

XX PD 21-OCT-1999.

XX PF 09-APR-1999; 99WO-IB00712.

XX PR 09-APR-1998; 98US-0057719.

XX PR 28-APR-1998; 98US-0069047.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-038446/03.

XX DR N-PSDB; AAZ42555.

XX PT Novel secreted protein 5' expressed sequence tag sequences used in
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
 XX Claim 3; Page 687; 837pp; English.

XX AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AAY64651 to
 CC AAY65438 represent the EST-related proteins corresponding to AAZ42265 to
 CC AAZ43052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be
 CC used for directing extracellular secretion of a polypeptide or the

CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AAZ42249 to AAZ42264 and AAY64644 to AAY64650 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 28 AA;

Query Match 18.5%; Score 40; DB 21; Length 28;
 Best Local Similarity 41.2%; Pred. No. 1.1e+02;
 Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAPGTFSTNTSTDICR 18
 ||||| :|||

Db 10 crpatldqatrattpcr 26

RESULT 7
 AAY24434
 ID AAY24434 standard; peptide; 26 AA.

XX AC AAY24434;

XX DT 27-SEP-1999 (first entry)

XX DE Adenovirus hexon protein heterologous ligand #2.

XX KW Adenovirus; RGD sequence; adenoviral vector; capsid protein; ligand;
 KW infection.

XX OS Mastadenovirus.

XX OS Synthetic.

XX PN WO9936545-A2.

XX PD 22-JUL-1999.

XX PF 15-JAN-1999; 99WO-US00913.

XX PR 16-JAN-1998; 98US-0071674.

XX PA (GENZ) GENZYME CORP.

XX PI Armentano D, O'Riordan CR, Romanczuk H;

XX DR WPI; 1999-444401/37.

XX PT Adenoviral vectors with modified capsid proteins for improved
 PT infectious capabilities

XX PS Example 1; Fig 1A; 59pp; English.

XX The present invention describes an adenoviral capsid protein comprising
 CC a heterologous ligand, where the ligand facilitates binding of the
 CC adenovirus to a target cell. The adenoviral vector is used to transfer a
 CC capsid protein (especially a fibre or hexon protein or protein IX)
 CC facilitates binding of the vector to the target cell. In particular, the
 CC adenoviral vector can be used to transfer the human cystic fibrosis
 CC transmembrane conductance regulator protein gene to the respiratory
 CC epithelium of test animals. The modified adenoviral capsid proteins
 CC improve and/or alter the infectious capability of the vector. The
 CC present sequence represents an adenovirus hexon protein heterologous
 CC ligand used in an example from the present invention.

XX SQ Sequence 26 AA;

Query Match 18.1%; Score 39; DB 20; Length 26;
 Best Local Similarity 47.4%; Pred. No. 1.3e+02;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 8 SNTTSSSTDICRPHQICNVV 26
 Db 1 sntssdqlaspyshprv 19

RESULT 8
 AAB44864
 ID AAB44864 standard; Protein: 34 AA.
 XX
 AC AAB44864;
 XX
 XX 09-FEB-2001 (first entry)
 XX Human secreted protein encoded by gene 35.
 DE Human; secreted protein; cytostatic; antiarthritic; antiasthmatic;
 XX immunosuppressive; antiarteriosclerotic; antiinflammatory; nootropic;
 KW neuroprotective; antidiabetic; tranquiliser; vulnerar; antibacterial;
 KW antipsoriatic; antiarrhythmic; antirheumatic; cardiac; anti-HIV;
 KW autoimmune disorder; allergic condition; cardiovascular disorder;
 KW cancer; neurological disease; tissue repair.
 XX
 OS Homo sapiens.
 XX
 XX WO200055176-A2.
 XX
 XX 21-SEP-2000.
 XX
 XX 09-MAR-2000; 2000WO-US06057.
 XX
 XX 12-MAR-1999; 99US-0124142.
 XX
 XX 11-JUN-1999; 99US-0138597.
 XX
 XX 03-DEC-1999; 99US-0168666.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI; 2000-638176/61.
 XX N-PSDB; AAC79883.
 XX
 XX Novel 49 human secreted proteins useful for diagnosis, prevention and
 PT treatment of disorders including neurological, cell proliferative,
 PT cardiovascular, and autoimmune/inflammatory disorders and microbial
 PT infections -
 XX
 XX Claim 11; Page 373-374; 405pp; English.
 XX
 XX This invention describes a novel isolated polypeptide (I) comprising an
 CC amino acid sequence at least 95 % identical to a polypeptide sequence
 CC selected from 49 polypeptides encoded by polynucleotide sequences
 CC included in American Type Culture Collection (ATCC) deposit number
 CC 203917, defined in the specification. The products of the invention have
 CC cytostatic, antiarthritic, antiasthmatic, immunosuppressive, nootropic,
 CC antiarteriosclerotic, antiinflammatory, neuroprotective, antidiabetic,
 CC tranquiliser, vulnerar, antibacterial, antipsoriatic, antiarrhythmic,
 CC antirheumatic, cardiac and anti-HIV activity. (I) or a nucleic acid (II)
 CC encoding (I) is useful for preventing, treating or ameliorating a medical
 CC condition and for diagnosing a pathological condition or susceptibility
 CC to the condition. (I) is useful for identifying a binding partner which
 CC affects the activity of the polypeptide and for identifying an activity
 CC in a biological sample. (I), (II) or an antibody (IV) specific to (I) is
 CC also useful for treating or preventing a disease, disorder or condition
 CC associated with aberrant expression of (I). Diseases treated or diagnosed
 CC include immune disorders such as autoimmune diseases, blood protein
 CC disorders, anemia, allergic reactions and conditions such as asthma,
 CC organ rejection or graft-versus-host disease, inflammation, hyper
 CC proliferative disorders, cardiovascular disorders such as arterioarterial
 CC fistula, arrhythmias, arteriosclerosis, coronary thrombosis, organ
 CC regeneration, cancer, neovascular glaucoma, diabetic retinopathy,
 CC rheumatoid arthritis, psoriasis, diseases associated with increased
 CC apoptosis that include acquired immunodeficiency syndrome (AIDS).

CC neurological diseases such as Parkinson's disease, viral, bacterial,
 CC fungal or parasitic diseases. They are also used to repair, replace or
 CC protect tissue damage by congenital defects, to treat trauma, in surgery,
 CC including cosmetic plastic surgery, to treat fibrosis, reperfusion injury
 CC or systemic cytokine damage, to stimulate chondrocyte growth, to prevent
 CC skin aging due to sunburn, to change a mammal's mental state or physical
 CC state by influencing biorhythms, cardiac rhythms, depression, memory,
 CC stress and for accelerating wound healing. (I), (II) and/or their agonist
 CC or antagonist are useful as food additives or preservatives to increase
 CC or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamin, mineral or other nutritional components. (I) is
 CC useful for screening therapeutic compounds. (II) is useful in forensic
 CC biology for detecting DNA sequences and as diagnostic probes for
 CC detecting the presence of specific mRNA in a particular cell type.

XX Sequence 34 AA;

Query Match 17.4%; Score 37.5; DB 21; Length 34;
 Best-Local Similarity 33.3%; Pred. No. 2.8e+02;
 Matches 6; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 6 TFSNTTSSSTDICRPHQIC 23

Db 13 tilnlatsttck-hdvc 29

RESULT 9

AAB51542

ID AAB51542 standard; Peptide: 22 AA.

XX

XX AAB51542;

XX 15-FEB-2001 (first entry)

XX Yada homologous peptide #5.

XX Proteobacteria; extracellular domain; virulence determinant; Yada;

KW adhesin; proteobacterial infection prevention; vaccine.

XX Thiobacillus ferrooxidans.

XX WO200061165-A1.

XX 19-OCT-2000.

XX 13-APR-2000; 2000WO-US09866.

XX 13-APR-1999; 99US-0129073.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX Lupas AN;

XX WPI; 2000-647397/62.

XX An isolated polypeptide conserved in proteobacterial extracellular

XX domains used in the treatment and prevention of bacterial infections -

XX Example 5; Page 59; 85pp; English.

XX This invention relates to peptides AAB51512 - AAB51537 which represent

XX conserved proteobacterial extracellular domains. Sequences

XX AAB51538 - AAB51618 represent peptides homologous to Yada, a yersinia

XX adhesin which is an important virulence determinant of the yersinia

XX species. The invention includes an antibody which binds to the

XX proteobacterial extracellular peptides, and an immunogenic composition

XX containing the antibody used as a vaccine to prevent infection by a

XX proteobacteria. The polypeptides and antibodies are useful in the

XX treatment and prevention of proteobacterial infections. The polypeptides

XX can also be used to identify compounds which antagonize the binding of a

XX bacterial adhesion to its ligand. The host cell can be used to produce

Best Local Similarity 46.7%; Pred. No. 2.8e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 15 DICRPHQICNVVAIP 29
|:|:|: | | | |
Db 6 dlcKpnsI--vveip 18

RESULT 12
AAW65481
ID AAY96558 standard; peptide; 22 AA.
XX AC AAY96558;
XX DT 12-SEP-2000 (first entry)
XX DE Human NF-ATc1 antigenic peptide 7.
XX KW NF-ATc1; cardiac hypertrophy; nuclear factor of activated T cells;
XX KW antagonist; congestive heart disease; cardiant; antigen.
XX OS Homo sapiens.
XX PN W0200030671-A2.
XX PD 02-JUN-2000.
XX PF 23-NOV-1999; 99W0-US27862.
XX PR 24-NOV-1998; 98US-0198977.
XX PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX PI Crabtree GR, Northrop JP, Ho SN;
XX WPI; 2000-399929/34.
XX PT Treating cardiac hypertrophy using NF-AT antagonists in
XX PT patients suffering from congestive heart disease
XX PS Disclosure; Page 28; 139pp; English.
XX CC A novel method for preventing and/or reducing cardiac hypertrophy in a
XX CC patient, comprises administering an NF-AT (nuclear factor of activated
XX CC T cells) antagonist to decrease the biological activity of NF-AT in
XX CC myocardial tissue (therefore preventing and/or reducing the level of
XX CC cardiac hypertrophy). The antagonists may decrease the transcriptional
XX CC activity, nuclear translocation or dephosphorylation of NF-AT, inhibit
XX CC binding of calcineurin to NF-AT, stimulate phosphorylation of NF-AT (e.g.
XX CC by increasing GSK-3) or inhibit formation of an NF-AT complex. The
XX CC antagonist is an antagonist of NF-ATc4 (also known as NF-AT3) and not
XX CC NF-ATc1, NF-ATc2 and NF-ATc3 (also known as NF-AT4). The method is used
XX CC for preventing and/or reducing cardiac hypertrophy in a patient suffering
XX CC from congestive heart disease (claimed) and for preventing other growth
XX CC of cardiac and vascular tissue.
XX SQ Sequence 22 AA;

Query Match 16.7%; Score 36; DB 21; Length 22;
Best Local Similarity 46.7%; Pred. No. 2.8e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 15 DICRPHQICNVVAIP 29
|:|:|: | | | |
Db 6 dlcKpnsI--vveip 18

RESULT 13
AAB66490
ID AAB66490 standard; Peptide; 22 AA.
XX AC AAB66490;

XX DT 10-APR-2001 (first entry)
XX DE Human NF-AT peptide #7.
XX KW Human; nuclear factor of activated T lymphocytes; NF-AT;
XX KW NF-AT cytoplasmic component; NF-ATc; NF-AT translocation;
XX KW nuclear localisation sequence; NLS; SRR.
XX OS Homo sapiens.
XX PN US6171781-B1.
XX PD 09-JAN-2001.
XX PF 27-MAR-1998; 98US-0049691.
XX PR 20-SEP-1993; 93US-0124981.
XX PR 13-JUN-1994; 94US-0260174.
XX PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX PI Crabtree GR, Northrop JP, Ho SN;
XX WPI; 2001-122328/13.
XX CC Screening assay for identifying modulators of translocation of nuclear
XX CC factor of activated T lymphocytes across nuclear membrane of cell, by
XX CC treating with nuclear factor of activated T lymphocyte polypeptide -
XX PS Disclosure; Column 23; 99pp; English.
XX CC The present sequence is given in a specification relating to a method for
XX CC identifying a compound which modulates translocation of a nuclear factor
XX CC of activated T lymphocytes (NF-AT) polypeptide across the nuclear
XX CC membrane of a cell. The method involves binding the compound to the
XX CC NF-AT polypeptide. The method is useful for identifying compounds which
XX CC modulate nuclear translocation of NF-AT present in the cytoplasm and/or
XX CC in the nucleus of the cell and for identifying agents that modulate
XX CC phosphorylation/dephosphorylation of NF-AT.
XX SQ Sequence 22 AA;

Query Match 16.7%; Score 36; DB 22; Length 22;
Best Local Similarity 46.7%; Pred. No. 2.8e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 15 DICRPHQICNVVAIP 29
|:|:|: | | | |
Db 6 dlcKpnsI--vveip 18

RESULT 14
AAW65481
ID AAW65481 standard; peptide; 23 AA.
XX AC AAW65481;
XX DT 12-OCT-1998 (first entry)
XX DE Hepatitis B surface antigen derived peptide (IGP 1082).
XX KW Annexin V; hepatitis B surface antigen; immunogen; vaccine;
XX KW hepatitis delta virus; infection; HBSAg.
XX OS Synthetic.
XX OS Hepatitis b virus.
XX PN W09829442-A1.
XX PD 09-JUL-1998.
XX XX

```

PF 23-DEC-1997; 97WO-EF07268.
XX
XX 11-JUL-1997; 97EP-0870103.
PR 30-DEC-1996; 96EP-0870164.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX De Meyer S, Depla E, Maertens G, Yap S;
XX WPI; 1998-388040/33.
XX
XX Immunogenic polypeptide from hepatitis B surface antigen - useful
PT in, e.g. vaccine against hepatitis B virus or hepatitis delta virus
PT infection
XX
XX Example 3; Page 35; 71pp; English.
XX
XX The invention relates to an immunogenic peptide derived from hepatitis B
CC surface antigen (HBsAg) which competes with the hepatitis B surface
CC antigen/annexin V interaction or which binds a compound or antibody
CC competing with the hepatitis B surface antigen/annexin V interaction.
CC Also claimed are: (1) a combination of the immunogenic peptide and a
CC negatively charged phospholipid; (2) a peptide composition comprising
CC the immunogenic peptide; (3) a vaccine comprising the immunogenic peptide
CC as an active substance; (4) antibodies which specifically bind to the
CC peptide and inhibit binding of HBsAg to annexin V, and (5) a therapeutic
CC composition comprising as an active substance the antibodies of (4).
CC The vaccine of (3), and the therapeutic composition of (5), can be used
CC as an inoculum to vaccinate humans against an infection with hepatitis
CC B and/or hepatitis delta virus. The immunogenic peptide can be used in a
CC method to detect antibodies which are capable of competing with the
CC hepatitis B and/or hepatitis Delta virus surface antigen/annexin V
CC interaction. The immunogenic peptide can also be used to screen for
CC drugs which block the binding between annexin V and the peptide, and as
CC a therapeutic to treat humans infected with hepatitis B virus and/or
CC hepatitis Delta virus. The present sequence represents one of the
CC peptide fragments derived from HBsAg which were synthesised to map the
XX annexin V-binding site on HBsAg.
XX
XX Sequence 23 AA;

```

```

Query Match 16.7%; Score 36; DB 19; Length 23;
Best Local Similarity 37.0%; Pred. No. 2.9e+02;
Matches 10; Conservative 4; Mismatches 7; Indels 6; Gaps 2;

```

```

QY 1 PCAPGTFSNTSSTICRPHQICNWA 27
Db | | | | : | | | | : | |
1 plipgt---sttsgpck---tctipa 21

```

```

RESULT 15
AAW97502
ID AAW97502 standard; peptide; 23 AA.
XX
XX AAW97502;
XX
XX 19-MAY-1999 (first entry)
XX
XX Antigenic site of HN protein loop beta-3L23.
XX
XX Antigenic site; haemagglutinin-neuraminidase; HN; paramyxoviridae;
KW virus epitope; attachment protein; vaccine; immunodominant epitope.
XX
XX Bovine parainfluenza virus.
XX
XX WO9902695-A2.
XX
XX 21-JAN-1999.
XX
XX 08-JUL-1998; 98WO-NL00390.
XX
XX 08-JUL-1997; 97EP-0202100.
PR

```

```

XX (DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.
PA Langedijk JPM, Van Oirschot JT;
XX
XX WPI; 1999-120896/10.
XX
XX Isolated proteinaceous substance - comprising at least one virus
PT epitope derived from an attachment protein of a paramyxovirus
PT
XX Disclosure; Page 46; 63pp; English.
XX
XX AAW97452-571 represent antigenic sites derived from the
CC haemagglutinin-neuraminidase (HN) protein of the paramyxoviridae.
CC The specification describes 3-D models identifying a proteinaceous
CC substance comprising at least one virus epitope derived from the
CC attachment protein, which corresponds to an antigenic site present on
CC one of the loops of HN. The antigenic sites can be used to produce
CC vaccines, to detect the viruses, and to select the immunodominant
CC epitope.
XX
XX Sequence 23 AA;

```

```

Query Match 16.7%; Score 36; DB 20; Length 23;
Best Local Similarity 42.9%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Qy 22 ICNVVAIPGNASMD 35
Db | | | | | | |
9 icnttgcpgkgtqr 22

```

```

Search completed: July 13, 2001, 17:13:20
Job time: 880 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:04:35 ; Search time 12.1 Seconds
(without alignments)
64.929 Million cell updates/sec

Title: US-09-800-909-2_COPY_163_201

Perfect score: 216

Sequence: 1 PCAPGTFSTSTDCRPHQICNVVAIFGNASMDAVCT 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 132220

Minimum DB seq length: 0

Maximum DB seq length: 39

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	100.0	39	1	US-08-050-319B-41
2	216	100.0	39	2	US-08-465-982-41
3	77	35.6	15	6	5395760-10
4	72	33.3	15	1	US-08-221-583-45
5	72	33.3	15	1	US-08-221-583-46
6	72	33.3	15	5	PCT-US95-04018-45
7	72	33.3	15	5	PCT-US95-04018-46
8	71	32.9	15	1	US-08-221-583-44
9	71	32.9	15	5	PCT-US95-04018-44
10	70	32.4	15	1	US-08-221-583-49
11	70	32.4	15	5	PCT-US95-04018-49
12	67	31.0	15	1	US-08-221-583-47
13	67	31.0	15	5	PCT-US95-04018-47
14	66	30.6	15	1	US-08-221-583-48
15	66	30.6	15	5	PCT-US95-04018-48
16	57	26.4	15	1	US-08-221-583-50
17	57	26.4	15	5	PCT-US95-04018-50
18	48	22.2	20	2	US-08-126-016-24
19	45	20.8	15	1	US-08-221-583-51
20	45	20.8	15	5	PCT-US95-04018-51
21	39	18.1	23	4	US-09-101-146-58
22	37.5	17.4	20	2	US-08-126-016-25
23	36	16.7	22	2	US-08-124-981A-28
24	36	16.7	22	3	US-09-037-190-31
25	36	16.7	22	3	US-09-037-192-31
26	36	16.7	22	4	US-09-037-143-31
27	36	16.7	22	4	US-09-049-691-31

28	36	16.7	22	4	US-08-260-174-31	Sequence 31, Appl
29	36	16.7	38	4	US-09-172-841-7	Sequence 7, Appl
30	35.5	16.4	36	4	US-08-944-483-26	Sequence 26, Appl
31	35.5	16.4	38	3	US-08-787-091-9	Sequence 9, Appl
32	35	16.2	27	1	US-08-599-556-7	Sequence 7, Appl
33	35	16.2	27	5	PCT-US96-05262-12	Sequence 12, Appl
34	35	16.2	38	1	US-08-444-005-18	Sequence 18, Appl
35	35	16.2	38	1	US-08-444-005-19	Sequence 19, Appl
36	34.5	16.0	30	1	US-08-428-248-1	Sequence 1, Appl
37	34.5	16.0	33	1	US-08-682-485A-24	Sequence 24, Appl
38	34.5	16.0	33	1	US-08-451-472-7	Sequence 7, Appl
39	34.5	16.0	33	2	US-08-933-314-24	Sequence 24, Appl
40	34	15.7	36	1	US-08-290-448A-23	Sequence 23, Appl
41	34	15.7	36	1	US-08-290-448A-23	Sequence 23, Appl
42	34	15.7	36	1	US-08-682-485A-7	Sequence 7, Appl
43	34	15.7	36	1	US-08-175-069A-23	Sequence 23, Appl
44	34	15.7	36	2	US-08-933-314-7	Sequence 7, Appl
45	34	15.7	37	1	US-08-682-485A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-050-319B-41
; Sequence 41, Application US/08050319B
; Patent No. 5633145
; GENERAL INFORMATION:
; APPLICANT: M.Feldmann, P.W. Gray,
; APPLICANT: M.J.C. Turner, F.M. Brennan
; TITLE OF INVENTION: Modified human TNFalpha (Tumor
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reed & Robbins
; STREET: 635 Bryant Street
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,319B
; FILING DATE: 10-May-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Robbins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5150-0030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-050-319B-41

Query Match 100.0%; Score 216; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PCAPGTFSTSTDCRPHQICNVVAIFGNASMDAVCT 39
|||||
DB 1 PCAPGTFSTSTDCRPHQICNVVAIFGNASMDAVCT 39
|||||

```
RESULT 2
US-08-465-982-41
; Sequence 41, Application US/08465982
; Patent No. 5863786
; GENERAL INFORMATION:
; APPLICANT: M.Feldmann, P.W. Gray,
; APPLICANT: M.J.C. Turner, F.M. Brennan
; TITLE OF INVENTION: Modified human TNFalpha (Tumor
; TITLE OF INVENTION: Necrosis Factor alpha) Receptor
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reed & Robbins
; STREET: 635 Bryant Street
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,982
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,319
; FILING DATE: 10-May-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Robbins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5150-0030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-982-41

Query Match 100.0%; Score 216; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNTSSDTCRPHQICNVVAIPGNASMDVACT 39
|||||
Db 1 PCAPGTFSTNTSSDTCRPHQICNVVAIPGNASMDVACT 39

RESULT 3
5395760-10
; Patent No. 5395760
; APPLICANT: SMITH, CRAIG A.; GOODWIN, RAYMOND G.; BECKMANN,
; M. PATRICIA
; TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND
; B-RECEPTORS
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/523,635
; FILING DATE: 10-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 421,417
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: 405,370
; FILING DATE: 11-SEP-1989
```

```
; APPLICATION NUMBER: 403,241
; FILING DATE: 05-SEP-1989
; SEQ ID NO:10;
; LENGTH: 15
5395760-10

Query Match 35.6%; Score 77; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 ICNVVAIPGNASMDA 36
|||||
Db 1 ICNVVAIPGNASMDA 15

RESULT 4
US-08-221-583-45
; Sequence 45, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavenr, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221.583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-221-583-45

Query Match 33.3%; Score 72; DB 1; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.0012;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GTFSNTTSSDTCRIP 19
|||||
Db 1 GTFSNTTSSDTCRIP 15

RESULT 5
US-08-221-583-46
; Sequence 46, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavenr, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25.mdtcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-46

Query Match 33.3%; Score 72; DB 1; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.0012;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 SNTSSTDICRPHQI 22
Db 1 SNTSSTDICRPHQI 15

RESULT 6
PCT-US95-04018-45
Sequence 45, Application PC/TUS9504018
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
APPLICANT: Kruszynski, Marian
APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: Norris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04018
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/221,580
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,583
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 46:
APPLICATION NUMBER: US 08/221,583

```

```

FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-45

Query Match 33.3%; Score 72; DB 5; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.0012;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GTFSNTSSTDICRP 19
Db 1 GTFSNTSSTDICRP 15

RESULT 7
PCT-US95-04018-46
Sequence 46, Application PC/TUS9504018
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
APPLICANT: Kruszynski, Marian
APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: Norris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04018
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/221,580
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,583
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 46:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-46

Query Match 33.3%; Score 72; DB 5; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.0012;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 SNTTSSTDICRPHOI 22
| | | | | | | | | | | | | | |
Db 1 SNTTSSTDICRPHOI 15

RESULT 8
US-08-221-583-44
; Sequence 44, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-44

Query Match 32.9%; Score 71; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APGTFSTNTSSTDI 16
| | | | | | | | | | | | | | |
Db 2 APGTFSTNTSSTDI 15

RESULT 9
PCT-US95-04018-44
; Sequence 44, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko

; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-44

Query Match 32.9%; Score 71; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APGTFSTNTSSTDI 16
| | | | | | | | | | | | | | |
Db 2 APGTFSTNTSSTDI 15

RESULT 10
US-08-221-583-49
; Sequence 49, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/221,583
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Deluca, Mark
;; REGISTRATION NUMBER: 33,229
;; REFERENCE/DOCKET NUMBER: CCOR-0185
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-3100
;; TELEFAX: (215) 568-3439
;; INFORMATION FOR SEQ ID NO: 49:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-221-583-49

Query Match 32.4%; Score 70; DB 1; Length 15;
Best Local Similarity 92.9%; Pred. No. 0.0022;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 RPHQICNVVAIPGN 31
||||| |||||||
DB 2 RPHQIANVVAIPGN 15

RESULT 11

PCT-US95-04018-49
;; Sequence 49, Application PC/TUS9504018
;; GENERAL INFORMATION:
;; APPLICANT: Heavner, George A.
;; APPLICANT: Kruszynski, Marian
;; APPLICANT: Mervic, Miljenko
;; APPLICANT: Weber, Robert W.
;; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
;; NUMBER OF SEQUENCES: 76
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
;; ADDRESSEE: Norris
;; STREET: One Liberty Place 46th Floor
;; CITY: Philadelphia
;; STATE: Pennsylvania
;; COUNTRY: USA
;; ZIP: 19403
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WordPerfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/04018
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/221,580
;; FILING DATE: 01-APR-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/221,583
;; FILING DATE: 01-APR-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/221,581
;; FILING DATE: 01-APR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Deluca, Mark
;; REGISTRATION NUMBER: 33,229
;; REFERENCE/DOCKET NUMBER: CCOR-0232
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-3100
;; TELEFAX: (215) 568-3439

;; INFORMATION FOR SEQ ID NO: 49:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US95-04018-49

Query Match 32.4%; Score 70; DB 5; Length 15;
Best Local Similarity 92.9%; Pred. No. 0.0022;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 RPHQICNVVAIPGN 31
||||| |||||||
DB 2 RPHQIANVVAIPGN 15

RESULT 12

US-08-221-583-47
;; Sequence 47, Application US/08221583
;; Patent No. 5486595
;; GENERAL INFORMATION:
;; APPLICANT: Heavner, George A.
;; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
;; NUMBER OF SEQUENCES: 62
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
;; STREET: One Liberty Place 46th Floor
;; CITY: Philadelphia
;; STATE: Pennsylvania
;; COUNTRY: USA
;; ZIP: 19403
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/221,583
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Deluca, Mark
;; REGISTRATION NUMBER: 33,229
;; REFERENCE/DOCKET NUMBER: CCOR-0185
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-3100
;; TELEFAX: (215) 568-3439
;; INFORMATION FOR SEQ ID NO: 47:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-221-583-47

Query Match 31.0%; Score 67; DB 1; Length 15;
Best Local Similarity 86.7%; Pred. No. 0.0056;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQICNV 25
||||| |||||||
DB 1 TSSTDICRPHQIANV 15

RESULT 13

PCT-US95-04018-47
;; Sequence 47, Application PC/TUS9504018
;; GENERAL INFORMATION:
;; APPLICANT: Heavner, George A.
;; APPLICANT: Kruszynski, Marian

APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: Norris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04018
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,580
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,583
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-47

Query Match 31.0%; Score 67; DB 5; Length 15;
Best Local Similarity 86.7%; Pred. No. 0.0056;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQICNV 25
||||| ||||| ||
Db 1 TSSTDICRPHQIANV 15

RESULT 14
US-08-221-583-48
Sequence 48, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25.mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-48

Query Match 30.6%; Score 66; DB 1; Length 15;
Best Local Similarity 86.7%; Pred. No. 0.0075;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 TDICRPHQICNVVAI 28
||||| ||||| |||||
Db 1 TDIARPHQIANVVAI 15

RESULT 15
PCT-US95-04018-48
Sequence 48, Application PC/TUS9504018
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
APPLICANT: Kruszynski, Marian
APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04018
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,580
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,583
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100

; TELFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-48

Query Match . 30.6%; Score 66; DB 5; Length 15;
Best Local Similarity 86.7%; Pred. No. 0.0075;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 14 TDICRPHQICNVVAI 28
 | | | | | | | | | |
Db 1 TDIARPHQIANVVAI 15

Search completed: July 13, 2001, 17:13:38
Job time: 543 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:38:33 ; Search time 13.26 Seconds
(without alignments) . 327.447 Million cell u

Title: US-09-800-909-2_COPY_201_257

Perfect score: 302
Sequence: 1 TSTSPTRSMAPGAVHLPQPV.....STSFLLPMGSPPPAEGSTGD 57

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:	13308
--	-------

```
Minimum DB seq length: 0
Maximum DB seq length: 57
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database :      PIR_68.*
1:  pir1.*
2:  pir2.*
3:  pir3.*
4:  pir4.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	71	23.5	54	2	A60726	mucin, gallbladder
2	56.5	18.7	52	2	S63324	hypothetical prote
3	55.5	18.4	42	2	I70082	glycoprotein Ib al
4	51	16.9	25	2	I40692	cena protein (IgA)
5	50	16.6	46	2	S07073	arabinogalactan pr
6	47	15.6	31	2	T36022	small hypothetical
7	44.5	14.7	57	2	S10782	salivary protein P
8	44.5	14.7	57	2	S1587	hypothetical prote
9	44	14.6	28	2	S6139	MHC class I HLA-J
10	43.5	14.4	47	2	S32108	sepiapterin reduct
11	43.5	14.4	54	2	D81737	hypothetical prote
12	43	14.2	31	2	I54515	pre-B cell Ig lamb
13	43	14.2	39	2	A45662	collagen alpha 2(V
14	42	13.9	36	2	A37172	collagen alpha 1(X
15	42	13.9	48	2	I45522	troponin T 2fa - r
16	42	13.9	55	2	S08424	H+-transporting AT
17	42	13.9	55	2	T11184	H+-transporting AT
18	41	13.6	42	2	T07030	extensin - tomato
19	40.5	13.4	50	2	A29789	mucin - sheep (fra
20	40	13.2	29	2	G36990	neural cell adhesi
21	40	13.2	50	2	H64801	hypothetical prote
22	40	13.2	51	2	A33756	dorsal protein - f
23	40	13.2	51	2	G72801	gp18 protein - Myc
24	40	13.2	52	2	S58216	hypothetical prote
25	40	13.2	54	2	T11131	ATP synthase subu
26	40	13.2	55	2	T11105	H+-transporting AT
27	40	13.2	57	2	I58120	gene insl protein
28	39.5	13.1	52	2	S01945	myosin catalytic l
29	39	12.9	27	2	S51176	aspartate transcar

30	39	12.9	53	2	S17672	hypothetical prote
31	38.5	12.7	51	2	C81117	hypothetical prote
32	38	12.6	53	2	S23202	kappa-casein - bov
33	38	12.6	54	2	S14338	serpin I - horse (
34	38	12.6	55	2	S29770	DNA-binding protei
35	38	12.6	55	2	A37238	autolimmune epitope
36	37.5	12.4	49	2	S25433	neural cell adhesi
37	37.5	12.4	51	2	T29481	hypothetical prote
38	37.5	12.4	57	2	E41715	hypothetical prote
39	37.5	12.4	57	2	T36648	hypothetical prote
40	37	12.3	30	2	JC1360	hypothetical 3K pr
41	37	12.3	38	2	S68260	hypothetical prote
42	37	12.3	45	2	S24713	Ig alpha chain, tr
43	37	12.3	49	2	S72213	beta-fructofuranos
44	37	12.3	51	2	B32040	dihydrodipsoamide S
45	37	12.3	55	2	A82795	hypothetical prote

ALIGNMENTS

```

RESULT      1
A60726      mucin, gallbladder - bovine (fragments)
C:Species:  Bos primigenius taurus (cattle)
C:Date:      28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993
C:Accession: A60726
R:Afdhal, N.H.; Offner, G.D.; Smith, B.F.
Gastroenterology 99, 1493-1501, 1990
A:Title:      Characterization of bovine gallbladder mucin. Amino acid sequences
A:Reference number: A60726; MUID:91007106
A:Accession:  A60726
A>Status:      preliminary
A:Molecule type: protein
A:Residues:    1-54 <AFD>

```

Query Match 23.5%; Score 71; DB 2; Length 54;
Best Local Similarity 35.2%; Pred. No. 2.3;
Matches 19; Conservative 6; Mismatches 25; Indels

QY 2 STSPTRMAPGAVHLPQPVSTRSQHT-QPTPEPSTAPSTSF---LLPMGPPSPA 51

pb 1 TTTTNTVTLPGSPQNTTTPSSLPETPNPTSTPPTVGGPSSPDPTSPGPPPS 54

RESULT 2

S63324
hypothetical protein YNL338w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N0170
C:Species: Saccharomyces cerevisiae
C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 05-Nov-1999
C:Accession: S63324
R:Obermaier, B.; Piravandi, E.; Rinke, M.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63317
A:Accession: S63324
A:Molecule type: DNA
A:Residues: 1-52 <OBE>
A:Cross-references: EMBL:Z71614; NID:gl302466; PID:e239576; PID:gl302467; G
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YNL338W
A:Map position: 14L

Query Match	18.7%	Score 56.5;	DB 2;	Length 52;
Best Local Similarity	35.7%	Pred. NO. 42;		
Matches 15: Conservative	4;	Mismatches 20;	Indels 3;	Gaps 2;

QY 8 SMAPGAVHLQPVPVSTRSQHTQTPEPSTAPSTSFLLPMGPSP 49
 || ::: | | | | | | | | | | | | | | | | | |
db 11 SMOYSDIYPPTPTPHHTTC--PTPHPH--PHTTHTTHHNPNP 49

A;Residues: 1-52 <OBE>
A:Cross-references: EMBL:Z71614: NID:Q1302466: PTD:Q1302467: GSPPDB:GN000

```

C:Keywords: hydroxyproline
F:6,7,9,11,14,16,18,20,26,30,34,36,39/Modified site: hydroxyproline (Pro) #status exp

Query Match      16.6%; Score 50; DB 2; Length 46;
Best Local Similarity 37.1%; Pred. No. 1.4e+02;
Matches 13; Conservative 4; Mismatches 16; Indels 2; Gaps 1;

QY 17 PQPVTSTRSQHTOPTPEPSTAPSTSFLLPMGPPSPA 51
      | | | | | | | | | | | | | | | | | | |
Db 6 PPAPAKPAPAPVPEASTAPVAA--PTTXPPSPA 38

RESULT 6
T36022
small hypothetical protein SCC54.09c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36022
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21581
A:Accession: T36022
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-31 <SEE>
A:Cross-references: EMBL:AL035591; PIDN:CAB38139.1; GSPDB:GN00070; SCODEB:GN00070; SCC54.09c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCODEB:SCC54.09c

Query Match      15.6%; Score 47; DB 2; Length 31;
Best Local Similarity 30.4%; Pred. No. 1.7e+02;
Matches 7; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 19 PVSTRSQHTOPTPEPSTAPSTSF 41
      | | | | | | | | | | | | | | |
Db 2 PLAARNEDNEPVPAPTWTGTGF 24

RESULT 7
SI0782
salivary protein P-B - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 19-Mar-1997 #sequence_revision 21-Nov-1998 #text_change 21-Nov-1998
C:Accession: SI0782
R:Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A:Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is a
A:Reference number: SI0780; MUID:90336641
A:Accession: SI0782
A:Molecule type: protein
A:Residues: 1-57 <STR>
C:Superfamily: proline-rich peptide P-B

Query Match      14.7%; Score 44.5; DB 2; Length 57;
Best Local Similarity 28.8%; Pred. No. 5.3e+02;
Matches 15; Conservative 2; Mismatches 28; Indels 7; Gaps 2;

QY 5 PTRSMAPGAVHLPQVPVSTRSQHTOPTPEPSTAPSTSFLLPMGPPSPAEGSTG 56
      | | | | | | | | | | | | | | | | | | |
Db 4 PRGPYPGGLAPPQFG--PGVPPPPPPYCGPGR-----IPPPPPAPYGP 48

RESULT 8
SI16587
hypothetical protein 1 - lamb's-quarters
C:Species: Chenopodium album (lamb's-quarters)
C>Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 26-May-1995
C:Accession: SI16587
R:Doerfel, P.; Weihe, A.; Dolferus, R.; Boerner, T.

C:Keywords: glycoprotein
A:Title: Polymorphism of human glycoprotein Ib alpha results from a variable number of
actions.
A:Reference number: I55355; MUID:92250564
A:Accession: I70082
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-42 <RES>
A:Cross-references: GB:S34439; NID:g249178; PIDN:AB22153.1; PID:g249179
C:Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein
C:Keywords: glycoprotein

Query Match      18.4%; Score 55.5; DB 2; Length 42;
Best Local Similarity 35.6%; Pred. No. 41;
Matches 16; Conservative 6; Mismatches 18; Indels 5; Gaps 3;

QY 5 PTRSMAPGAVHLPQVPVSTRSQHTOPTPEPSTAPSTSFLLPMGPPSP 49
      | | | | | | | | | | | | | | | | | | |
Db 2 PTSEPAESPPT-TPEPTEPAP-SPTTPEPTEPAPS---PTTPEP 41

RESULT 4
I40692
cenaA protein (IgaIh) - Cellulomonas fimi (fragment)
C:Species: Cellulomonas fimi
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 22-Oct-1999
C:Accession: I40692
R:Miller, P.B.; Shen, H.; Gilkes, N.R.; Kilburn, D.G.; Miller, R.C.; Plaut, A.G.; Warren
FEMS Microbiol. Lett. 92, 199-204, 1992
A:Title: Endoglucanase A from Cellulomonas fimi in which the hinge sequence of human IgA
hoaae.
A:Reference number: I40692
A:Accession: I40692
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25 <RES>
A:Cross-references: EMBL:X65780; NID:g312035; PIDN:CAA46663.1; PID:g312036

Query Match      16.9%; Score 51; DB 2; Length 25;
Best Local Similarity 46.2%; Pred. No. 60;
Matches 12; Conservative 1; Mismatches 5; Indels 8; Gaps 1;

QY 24 SQHTQPTPEPSTAPSTSFLLPMGPPSP 49
      | | | | | | | | | | | | | | |
Db 2 SVSTPTPTSPSTPT-----PSP 19

RESULT 5
S07073
arabinogalactan protein - Italian ryegrass (fragments)
C:Species: Lolium multiflorum (Italian ryegrass)
C>Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 05-Dec-1998
C:Accession: S07073
R:Gleeson, P.A.; McNamara, M.; Wattenhall, R.E.H.; Stone, B.A.; Fincher, G.B.
Biochem. J. 264, 857-862, 1989
A:Title: Characterization of the hydroxyproline-rich protein core of an arabinogalactan-
A:Reference number: S07073; MUID:90147544
A:Accession: S07073
A>Status: preliminary
A:Molecule type: protein
A:Residues: 446 <GLE>
A:Note: 19-His, and 23-Leu were also found
```

[illegible]

R; Mayne, R.; Brewton, R.G.; Mayne, P.M.; Baker, J.R.

J. Biol. Chem. 268, 9381-9386, 1993

A; Title: Isolation and characterization of the chains of type V/type XI collagen present

A; Reference number: A46662; MUID: 93252802

A; Accession: A46662

A; Status: preliminary

A; Molecule type: protein

A; Residues: 1-39 <MAY>

A; Experimental source: vitreous humor

A; Note: sequence extracted from NCBI backbone (NCBIP:131547)

C; Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Search completed: July 13, 2001, 17:41:18

Job time: 165 sec

Query Match 14.2%; Score 43; DB 2; Length 39;
Best Local Similarity 53.3%; Pred. No. 4.8e+02;
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 43 LPMGPPPPAEGSTGD 57

|| || || || ||

Db 14 LPGPPPPGEAGPGD 28

RESULT 14

A37172

collagen alpha 1(XII) chain-like, skin and tendon - bovine (fragments)

C; Species: Bos primigenius taurus (cattle)

C; Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 19-Oct-1995

C; Accession: A37172

R; Dublet, B.; Van Der Rest, M.

Ann. N. Y. Acad. Sci. 580, 436-439, 1989

A; Title: Comparison between chicken type XII collagen and bovine homologues.

A; Reference number: A37172

A; Accession: A37172

A; Status: preliminary

A; Molecule type: protein

A; Residues: 1-36 <DUB>

Query Match 13.9%; Score 42; DB 2; Length 36;
Best Local Similarity 41.4%; Pred. No. 5.4e+02;
Matches 12; Conservative 3; Mismatches 8; Indels 6; Gaps 3;

QY 25 QHTQTPPESTA-PSTSFLLPMGPPPAE 52

| | | | | | | | | | | | | |

Db 3 QERSP-PQANAVESX---PASPSPLQ 26

RESULT 15

I46522

troponin T 2fa - rabbit (fragment)

C; Species: Oryctolagus cuniculus (domestic rabbit)

C; Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999

C; Accession: I46522

R; Briggs, M.M.; Lin, J.J.; Schachat, F.H.

J. Muscle Res. Cell. Motil. 8, 1-12, 1987

A; Title: The extent of amino-terminal heterogeneity in rabbit fast skeletal muscle tropo

A; Reference number: I46522; MUID: 87251333

A; Accession: I46522

A; Status: preliminary; translated from GB/EMBL/DDBJ

A; Molecule type: mRNA

A; Residues: 1-48 <BRI>

A; Cross-references: EMBL:U04975; NID:9440810; PID:AAA16028.1; PID:9440811

C; Superfamily: troponin T

Query Match 13.9%; Score 42; DB 2; Length 48;
Best Local Similarity 37.9%; Pred. No. 7.3e+02;
Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 2;

QY 10 AGAVHLPPQVSTRSQHTQTPPEPS-TAP 37

| | | | | | | | | | | | | |

Db 25 SPAGVHEPEV-----HEEEKPRPKLTAP 48

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:41:23 ; Search time 9.89 Seconds
(without alignments)
197.428 Million cell updates/sec

Title: us-09-800-909-2_COPY_201_257

Perfect score: 302
Sequence: 1 TSTSPRSMAPGAVHLPQPV.....STSFLLPMGPSPPAEGSTGD 57

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 4271

Minimum DB seq length: 0
Maximum DB seq length: 57

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	56.5	18.7	52	1 YN78_YEAST	P53820 saccharomyc
2	44.5	14.7	57	1 PRPB_HUMAN	P02814 homo sapien
3	42	13.9	55	1 ATP8_GADMO	P15996 gadus morhu
4	40	13.2	51	1 VG18_BPMD2	O64211 mycobacteri
5	40	13.2	55	1 ATP8_PELSU	O79674 pelomedusa
6	39	12.9	56	1 T46P_HUMAN	O923f1 homo sapien
7	38	12.6	51	1 MLEV_MOUSE	P09542 mus musculu
8	38	12.6	55	1 ATP8_PAROL	O959d5 paralichthy
9	37.5	12.4	16	1 FOR2_MYRGU	P81437 myrmecia gu
10	37.5	12.4	57	1 RPOK_HALMA	P29200 halocarcula
11	36.5	12.1	55	1 YPU3_RHOCA	P26159 rhodobacter
12	36	11.9	55	1 ATP8_SQUAC	O92250 squalus aca
13	35	11.6	53	1 CAL1_RABIT	P02456 oryctolagus
14	35	11.6	55	1 ATP8_SALAL	O9xn27 salvelinus
15	34.5	11.4	16	1 FOR1_MYRGU	P81438 myrmecia gu
16	34	11.3	55	1 ATP8_RHEAM	O79396 rhea ameri
17	34	11.3	55	1 M84C_DROME	O01644 drosophila
18	33.5	11.1	38	1 H5_COLLI	P02260 columba liv
19	33.5	11.1	52	1 MTK_DROME	O24395 drosophila
20	33.5	11.1	52	1 PH68_HUMAN	O9xht9 homo sapien
21	33.5	11.1	55	1 ATP8_AYTAM	O9xxkz5 aethya amer
22	33	10.9	32	1 CAPP_METEX	O49136 methylobact
23	33	10.9	33	1 PBAN_LYMDI	P43511 lymantria d
24	33	10.9	50	1 ZMT4_BOVIN	O9rtf3 bos taurus
25	33	10.9	51	1 TAT_HV1J3	P12508 human immu
26	33	10.9	54	1 ATP8_CARAU	O78683 carassius a
27	33	10.9	54	1 ATP8_CYPCA	P24948 cyprinus ca
28	33	10.9	55	1 ATP8_LOXNO	O9mdj1 loxigilla n
29	32	10.6	34	1 RNLL_PIG	P15466 sus scrofa
30	32	10.6	50	1 SPRT_RAT	P81728 rattus norv
31	32	10.6	55	1 ATP8_LATCH	O03168 latimeria c
32	32	10.6	55	1 ATP8_SALFO	O9xn35 salvelinus
33	31.5	10.4	17	1 A45K_MYCBO	P80069 mycobacteri

34 31.5 10.4 55 1 ATP8_STRCA O21401 struthio ca
35 31 10.3 20 1 DFTS_RAT P07448 rattus norv
36 31 10.3 20 1 UCRQ_EQUAR P81247 equisetum a
37 31 10.3 46 1 YPC4_ECOLI P19755 escherichia
38 31 10.3 55 1 ATP8_CROLA P34190 crossostoma
39 30.5 10.1 50 1 PEN1_PENVA P81056 penaeus van
40 30.5 10.1 52 1 RUBR_DESVH P00269 desulfovibr
41 30.5 10.1 55 1 A70A_DROSE O18417 drosophila
42 30 9.9 26 1 CATG_RAT P17977 rattus norv
43 30 9.9 34 1 EML_ENSMI P27205 ensis minor
44 30 9.9 34 1 H1S_STRPU P19376 strongyloce
45 30 9.9 36 1 NUCM_SOLTU P80264 solanum tub

ALIGNMENTS

RESULT 1
IN78_YEAST STANDARD; PRT; 52 AA.
AC P53820;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 6.0 KDA PROTEIN IN COSI 5 REGION.
GN YNL338W OR N0170.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rink M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YHR217C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z71614; CAA96274.1; -;
DR EMBL; Z71613; CAA96273.1; -;
DR SGD; S0005282; YNL338W.
KW Hypothetical protein.
SQ SEQUENCE 52 AA; 5951 MW; C1E4066D43E057A1 CRC64;

Query Match 18.7%; Score 56.5; DB 1; Length 52;
Best Local Similarity 35.7%; Pred. No. 30;
Matches 15; Conservative 4; Mismatches 20; Indels 3; Gaps 2;

QY 8 SMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTSFLLPMGPSP 49
|| ::||| ||||| |||||
Db 11 SMQYSDIIIPPTPTHTHTT-PTPHPH--PHTHTHTHHNP 49

RESULT 2

ID PRPB_HUMAN STANDARD; PRT; 57 AA.
AC P02814;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROLINE-RICH PEPTIDE P-B [CONTAINS: PEPTIDE P-A].
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]

```

DR InterPro: IPR001421; -
DR Pfam: PF00695; ATP-synt_8; 1.
DR KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
DR TRANSMEM 4 24 POTENTIAL.
DR FT SEQUENCE 55 AA; 6481 MW; E85C81E63DB48B15 CRC64;
DR SQ SEQUENCE 55 AA; 6481 MW; E85C81E63DB48B15 CRC64;

Query Match 13.9%; Score 42; DB 1; Length 55;
Best Local Similarity 32.0%; Pred. No. 5.3e+02;
Matches 8; Conservative 5; Mismatches 12; Indels 0; Gaps 1;

QY 16 LPPQVSTRSOHTQTPPEPSTAPSTS 40
   || | : :||: ||:
DB 24 LPPKVAHTFPNEPSPQGMTPKTA 48

RESULT 4
VG18_BPMD2 STANDARD; PRT; 51 AA.
ID VG18_BPMD2 STANDARD; PRT; 51 AA.
AC O64211;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENE 18 PROTEIN (GP18).
DE 18.
GN OS Mycobacteriophage D29.
OS OC Mycobacteriophage D29.
OC Viruses.
OX NCBI_TaxID=28369;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: implications for phage
   evolution.";
RT J. Mol. Biol. 279:143-164(1998).
CC -----CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC -----CC
DR EMBL: AF022214; AAC18458.1; -
DR SEQUENCE 51 AA; 5680 MW; EF85B1AFF5786A34 CRC64;

Query Match 13.2%; Score 40; DB 1; Length 51;
Best Local Similarity 33.3%; Pred. No. 7.3e+02;
Matches 11; Conservative 4; Mismatches 14; Indels 4; Gaps 1;

QY 7 RSMAGAVHLPPQVSTRSOHTQTPPEPSTAPST 39
   | | | | | : | | | | | | | |
DB 21 RLTIAGGWRPRKPRPRT---TKPKPAPKQEPAT 49

RESULT 5-
ATP8_PELSU STANDARD; PRT; 55 AA.
ID ATP8_PELSU STANDARD; PRT; 55 AA.
AC O79674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
DE GN MTATP8 OR ATP8.
OS Pelomedusa subrufa (African side-necked turtle).
OG Mitochondrion.
OG Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Pleurodira; Pelomedusidae; Pelomedusa.
OX NCBI_TaxID=44522;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Zardoya R.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
 CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
 CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF039066; AAD05054.1; -
 DR InterPro; IPR001421; -
 DR Pfam; PF00895; ATP-synt_8; 1.
 KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
 FT TRANSMEM 4 24 POTENTIAL.
 SQ SEQUENCE 55 AA; 6536 MW; D8D4BC8F8651A001 CRC64;

 Query Match 13.2%; Score 40; DB 1; Length 55;
 Best Local Similarity 36.0%; Pred. No. 7.8e+02;
 Matches 9; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
 QY 15 HLPQPVSTRSQHTQTPETSTAPST 39
 DB 31 HPPNSPKNNKLTPHPWTWPT 55

 RESULT 6
 TA6P_HUMAN
 ID TA6P_HUMAN STANDARD; PRT; 56 AA.
 AC Q9X3F1.
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TAP2-ASSOCIATED 6.5 KDA POLYPEPTIDE.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Li H.C.;
 RT "A 56 aa polypeptide with phosphorylation motif, potentially
 RT associated with Tap2 isoform activity."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE ASSOCIATED WITH TAP2 ISOFORM ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF152583; AAD32715.1; -
 SQ SEQUENCE 56 AA; 6535 MW; ACD5D223ECC3C3BA CRC64;

Query Match 12.9%; Score 39; DB 1; Length 56;
 Best Local Similarity 40.0%; Pred. No. 9.6e+02;
 Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 30 TPEPSTAPSTSLPMPGSP 49
 DB 6 TPQILTISFVSILSPSP 25

RESULT 7
 MLEV_MOUSE
 ID MLEV_MOUSE STANDARD; PRT; 51 AA.
 AC P09542;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE MYOSIN LIGHT CHAIN 1, SLOW-TWITCH MUSCLE B/VENTRICULAR ISOFORM
 DE (FRAGMENT).
 GN MYL3 OR MLC1V.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H; TISSUE=Spleen;
 RX MEDLINE=89057447; PubMed=3194193;
 RA Cohen A., Barton P.J.R., Robert B., Garner I., Alonso S.,
 RA Buckingham M.E.;
 RT "Promoter analysis of myosin alkali light chain genes expressed in
 RT mouse striated muscle."
 RL Nucleic Acids Res 16:10037-10052(1988).
 CC -!- SUBUNIT: MYOSIN IS AN HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
 CC PROTEIN DOES NOT BIND CALCIUM.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X12972; CAA31415.1; -
 DR FIR; S01945; S01945.
 DR HSSP; P04002; IATE.
 KW Myosin; Muscle protein; Multigene family.
 FT INIT_MET 0
 FT NON_TER 51 51
 SQ SEQUENCE 51 AA; 5085 MW; CE513EBA3C8258D CRC64;

 Query Match 12.6%; Score 38; DB 1; Length 51;
 Best Local Similarity 26.3%; Pred. No. 1.1e+03;
 Matches 10; Conservative 6; Mismatches 20; Indels 2; Gaps 1;
 QY 17 PQVSTRSQHTQTPETSTAPSTSLPMPGSPPAE 52
 DB 5 PEKDDAKAAPAPAAAPAAAPAAAPAAPEPKE 42

 RESULT 8
 ATP8_PAROL
 ID ATP8_PAROL STANDARD; PRT; 55 AA.
 AC Q9T9P5;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
 GN MTATP8 OR ATP8.
 OS Paralichthys olivaceus (Flounder).
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 CC Pleuronectoidel; Bothidae; Paralichthys.
 OX NCBI_TaxID=8255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Saitoh K., Hayashizaki K., Yokoyama Y., Asahida T., Toyohara H.,
 RA Yamashita Y.;

RT "The complete nucleotide sequence of Japanese flounder mitochondrial
 genome: structural property and cue for resolving teleostean
 relationship."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
 CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB028664; BAA89037.1; -
 DR InterPro; IPR001421; -
 DR Pfam; PF00895; ATP-synt_8; 1.
 KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
 SQ SEQUENCE 55 AA; 6571 MW; 2B5FEF20FDCB6AA9 CRC64;

Query Match 12.6%; Score 38; DB 1; Length 55;
 Best Local Similarity 29.2%; Pred. No. 1.1e+03;
 Matches 7; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 16 LPQVSTRSQHTQTPPESTAPST 39
 :| | :| | :| | :| |
 DB 24 IPPKVLATFPNEPTPOSTQPKT 47

RESULT 9

FOR2_MYRGU STANDARD; PRT; 16 AA.
 AC P81437;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE FORMARCIN 2.
 OS Myrmecia gulosa (Red bulldog ant).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 OC Formicoidea; Formicidae; Myrmecia.
 OX NCBI_TaxID=36170;
 RN [1]
 RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
 RC TISSUE=Hemolymph;
 RX MEDLINE=98165787; PubMed=9497332;
 RA Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
 RT "Isolation from an ant Myrmecia gulosa of two inducible
 O-glycosylated proline-rich antibacterial peptides."
 RL J. Biol. Chem. 273:6139-6143(1998).
 CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE. HAS ACTIVITY AGAINST E.COLI
 CC BUT NONE AGAINST OTHER GRAM-NEGATIVE BACTERIA AND GRAM-POSITIVE
 CC BACTERIA.
 CC -1- INDUCTION: UPON BACTERIAL CHALLENGE.
 CC -1- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
 CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
 CC -1- SIMILARITY: TO DROSOPHILA DROSOCIN.
 KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
 FT CARBOHYD 11
 FT CARBOHYD 11 O-LINKED (GALNAC-);
 SQ SEQUENCE 16 AA; 1807 MW; 9C3CA3B00BC2E0AE CRC64;

Query Match 12.4%; Score 37.5; DB 1; Length 16;
 Best Local Similarity 47.1%; Pred. No. 3.9e+02;
 Matches 8; Conservative 2; Mismatches 2; Indels 5; Gaps 1;

QY 17 PQPVSTRSQHTQTPPEP 33
 | | | | |

Db 3 PNPVNTK-----PTYP 14

RESULT 10

RPOK_HALMA STANDARD; PRT; 57 AA.
 AC P29200;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE SUBUNIT K (EC 2.7.7.6).
 GN RPOK.
 OS Haloarcula marismortui (Halo bacterium marismortui).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloarcula.
 OX NCBI_TaxID=2238;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92105119; PubMed=1840597;
 RA Kroemer W.J., Arndt E.;
 RT "Halo bacterium S9 operon. Three ribosomal protein genes are
 RT cotranscribed with genes encoding a tRNA(Leu), the enolase, and a
 RT putative membrane protein in the archaeobacterium Haloarcula
 RT (Halo bacterium) marismortui."
 RL J. Biol. Chem. 266:24573-24579(1991).
 RN [2]
 RP SIMILARITY.
 RX MEDLINE=94321350; PubMed=8045907;
 RA McKune K., Woychik N.A.;
 RT "Halo bacterium S9 operon contains two genes encoding proteins
 RT homologous to subunits shared by eukaryotic RNA polymerases I, II,
 RT and III."
 RL J. Bacteriol. 176:4754-4756(1994).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +
 CC RNA(N).
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOK / EUKARYOTIC RPB6
 CC RNA POLYMERASE SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M76567; AAA73100.1; -
 DR PIR; E41715; E41715.
 DR InterPro; IPR001725; -
 DR Pfam; PF01192; RNA_pol_K; 1.
 DR PROSITE; PS01111; RNA_POL_K_14KD; 1.
 KW Transferase; DNA-directed RNA polymerase; Transcription.
 SQ SEQUENCE 57 AA; 6291 MW; D299FDA1954D030F CRC64;

Query Match 12.4%; Score 37.5; DB 1; Length 57;
 Best Local Similarity 42.9%; Pred. No. 1.3e+03;
 Matches 9; Conservative 3; Mismatches 4; Indels 5; Gaps 1;

QY 9 MAPGAVHLPQPVSTRSQHTQ 29
 :| | | | :| | | |
 DB 21 LAHGA-----PVLIEHTQ 36

RESULT 11

YPU3_RHOCA STANDARD; PRT; 55 AA.
 AC P26159;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-MAY-1992 (Rel. 22, Last annotation update)

```

Query Match      11.9%; Score 36; DB 1; Length 55;
Best Local Similarity 35.3%; Pred. No. 1.7e+03;
Matches 12; Conservative 2; Mismatches 12; Indels 8; Gaps 1;

QY 14 VHLPPQVSTRSQHTQPT-----PEPSTAPST 39
      + + + + + : + + + + + + + + + + +
DB 22 VILPKKVMTHLFNNPTAKSAEKPKPEPWNWPT 55

RESULT 13
CALL_RABIT
ID CALL_RABIT STANDARD; PRT; 53 AA.
AC P02456;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE COLLAGEN ALPHA 1(I) CHAIN (FRAGMENT).
GN COL1A1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1].
RP SEQUENCE.
RX MEDLINE=70252720; PubMed=4194291;
RA Bornstein P., Nesse R.;
RT "The comparative biochemistry of collagen: the structure of rabbit
RT skin collagen and its relevance to immunochemical studies of
RT collagen.";
RL Arch. Biochem. Biophys. 138:443-450(1970).
CC -I- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC (FIBRILLAR FORMING COLLAGEN).
CC -I- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
CC -I- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
CC HYDROXYAPATITE.
CC -I- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
DR PIR; A02856; CGBL5.
DR InterPro; IPR001007; -.
DR PROSITE; PS01208; VWFC; PARTIAL.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Collagen.
FT FT MOD_RES 7 7 CONVERTED TO AN ALDEHYDE GROUP THAT IS
FT FT MOD_RES 26 26 INVOLVED IN CROSS-LINKING.
FT FT MOD_RES 29 29 HYDROXYLATION (PROBABLE).
FT FT MOD_RES 32 32 HYDROXYLATION (PROBABLE).
FT FT MOD_RES 41 41 HYDROXYLATION (PROBABLE).
FT FT MOD_RES 44 44 HYDROXYLATION (PROBABLE).
FT FT MOD_RES 47 47 HYDROXYLATION (PROBABLE).
FT FT NON_TER 53 53
SQ SEQUENCE 53 AA; 4987 MW; 127582E5E52B87FC CRC64;

Query Match      11.6%; Score 35; DB 1; Length 53;
Best Local Similarity 30.2%; Pred. No. 2e+03;
Matches 13; Conservative 4; Mismatches 24; Indels 2; Gaps 2;

QY 14 VHLPPQVSTRSQHTQPTTPPSTPSTFSLPMGPPSPAEGSTG 56
      + + + + + + + + + + + + + + + + +
DB 11 VSVPGMGSGRGLPGPGAPCPZ-GFZGPPG-2PGZPGSSG 51

RESULT 14
ATP8_SALAL
ID ATP8_SALAL STANDARD; PRT; 55 AA.
AC Q9XN27;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)

```

DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
GN MTATP8 OR ATP8 OR ATPASE8.
OS Salvelinus alpinus (Arctic char).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8036;
RN [1]
RP SEQUENCE FROM N.A.
RA Dolron S., Blier P.U., Bernatchez L.;
RT "A comparative analysis of complete sequence of mitochondrial genome
RT between brook char (Salvelinus fontinalis) and arctic char (S.
RT alpinus).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF154851; AAD41389.1; -.
DR InterPro; IPR001421; -.
DR Pfam; PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
KW SEQUENCE 55 AA; 6455 MW; 71E430C2E346924A CRC64;
SQ

Query Match 11.6%; Score 35; DB 1; Length 55;
Best Local Similarity 30.0%; Pred. No. 2e+03;
Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 12 GAVHLPQVSTRSQHTQTP 31
| : : | : : | : : |
Db 30 GHFTNEPTSTOSTKTP 49

RESULT 15
FORL_MYRGU
ID FORL_MYRGU STANDARD; PRT; 16 AA.
AC P81438;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FORMACIN 1.
OS Myrmecia gulosa (Red bulldog ant).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Formicoidea; Formicidae; Myrmecia.
OX NCBI_TaxID=36170;
RN [1]
RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
RC TISSUE=Hemolymph;
RX MEDLINE=98165787; PubMed=9497332;
RA Mackintosh J.A., Veal D.A., Beattie A.J., Cooley A.A.;
RT "Isolation from an ant Myrmecia gulosa of two inducible
RT O-glycosylated proline-rich antibacterial peptides.";
RL J. Biol. Chem. 273:6139-6143(1998).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE. HAS ACTIVITY AGAINST E.COLI
CC BUT NONE AGAINST OTHER GRAM-NEGATIVE BACTERIA AND GRAM-POSITIVE
CC BACTERIA.
CC -!- INDUCTION: UPON BACTERIAL CHALLENGE.
CC -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.

FT CARBOHYD 11 11 O-LINKED (GALNAC...).
SQ SEQUENCE 16 AA; 1794 MW; 80CEA3AABBC2E0AE CRC64;

Query Match 11.4%; Score 34.5; DB 1; Length 16;
Best Local Similarity 41.2%; Pred. No. 7e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 5; Gaps 1;

Qy 17 POPVSTRSQHTQTP 33
| : : | : : | : : |
Db 3 PNPVANK-----PTPHP 14

Search completed: July 13, 2001, 17:44:04
Job time: 161 sec


```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92250564; PubMed=1577776;
RA Lopez J.A., Ludwig E.H., McCarthy B.J.;
RT "Polymorphism of human glycoprotein Ib alpha results from a variable
number of tandem repeats of a 13-amino acid sequence in the mucin-like
RT macroglycopeptide region. Structure/function implications.";
RL J. Biol. Chem. 267:10055-10061(1992).
DR EMBL; S34439; AAB22153.1; -.
FT NON_TER 1
SQ SEQUENCE 42 AA; 4222 MW; 228018AC7FBE3F38 CRC64;

Query Match 18.4%; Score 55.5; DB 4; Length 42;
Best Local Similarity 35.6%; Pred. No. 23;
Matches 16; Conservative 6; Mismatches 18; Indels 5; Gaps 3;

QY 5 PTRSMAGAVHLPQVSTRSQHTQTPPEPSTAPSTSFLLPMGPSP 49
II II : : : : : : : : : : : : : : : :
Db 2 PTSEPA SPT-TPEPTSEPA SPT-TPEPTSEPA SPT-TPEPTSEPA 41

RESULT 3
O92313 ID O92313 PRELIMINARY; PRT; 52 AA.
AC O92313;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=196775;
RA Coggin W.B., Lefkowitz E.J., Sullender W.M.;
RT "Genetic Variability among Group A and Group B Respiratory Syncytial
RT Viruses in a Children's Hospital.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF086886; AAC43006.1; -.
DR InterPro; IPR000925; -.
DR Pfam; PF00802; Glycoprotein_G; 1.
FT NON_TER 1
FT NON_TER 52
SQ SEQUENCE 52 AA; 5541 MW; 8C3111FE05DF29E0 CRC64;

Query Match 17.2%; Score 52; DB 14; Length 52;
Best Local Similarity 34.3%; Pred. No. 65;
Matches 12; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 15 HLPQPVSTRSQHTQTPPEPSTAPSTSFLLPMGPSP 49
II II : : : : : : : : : : : : : : : :
Db 13 HTSQKETHLSTSEGNPSPSQVYTTSEYLSQSPSP 47

RESULT 4
O92301 ID O92301 PRELIMINARY; PRT; 52 AA.
AC O92301;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
```

```
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=181691;
RA Coggin W.B., Lefkowitz E.J., Sullender W.M.;
RT "Genetic Variability among Group A and Group B Respiratory Syncytial
RT Viruses in a Children's Hospital.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF086872; AAC42992.1; -.
DR InterPro; IPR000925; -.
DR Pfam; PF00802; Glycoprotein_G; 1.
FT NON_TER 1
FT NON_TER 52
SQ SEQUENCE 52 AA; 5613 MW; 7C31114ACA02574E6 CRC64;

Query Match 16.9%; Score 51; DB 14; Length 52;
Best Local Similarity 34.3%; Pred. No. 82;
Matches 12; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 15 HLPQPVSTRSQHTQTPPEPSTAPSTSFLLPMGPSP 49
II II : : : : : : : : : : : : : : : :
Db 13 HTSQKETHLSTSEGNPSPSQVYTTSEYLSQSPSP 47

RESULT 5
O92314 ID O92314 PRELIMINARY; PRT; 52 AA.
AC O92314;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=203721;
RA Coggin W.B., Lefkowitz E.J., Sullender W.M.;
RT "Genetic Variability among Group A and Group B Respiratory Syncytial
RT Viruses in a Children's Hospital.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF086888; AAC43008.1; -.
DR InterPro; IPR000925; -.
DR Pfam; PF00802; Glycoprotein_G; 1.
FT NON_TER 1
FT NON_TER 52
SQ SEQUENCE 52 AA; 5554 MW; 3B328FBC5B4E9858 CRC64;

Query Match 16.9%; Score 51; DB 14; Length 52;
Best Local Similarity 31.8%; Pred. No. 82;
Matches 14; Conservative 5; Mismatches 25; Indels 0; Gaps 0;

QY 6 TRSMAGAVHLPQVSTRSQHTQTPPEPSTAPSTSFLLPMGPSP 49
II II : : : : : : : : : : : : : : : :
Db 4 TNSITGNLEHTSQEETHLSTSEGNTPSQAYTTSEYLSQSPSP 47

RESULT 6
O02832 ID O02832 PRELIMINARY; PRT; 33 AA.
AC O02832;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE HUNTINGTIN PROTEIN (FRAGMENT).
GN IT15.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
```

```

OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96326790; PubMed=8766138;
RX Pecheux C., Gall A.L., Kaplan J.C., Dode C.;
RT "Sequence analysis of the CAG triplet repeats region in the Huntington
RL disease gene (It15) in several mammalian species.";
RD Ann. Genet. 39:81-86(1996).
DR ENBL; S83377; AAB50771.1; -.
FT NON_TER 1
SQ SEQUENCE 33 AA; 3393 MW; 44F558CE5636722A CRC64;

Query Match 16.7%; Score 50.5; DB 6; Length 33;
Best Local Similarity 38.5%; Pred. No. 62;
Matches 15; Conservative 1; Mismatches 12; Indels 11; Gaps 3;

QY 16 LPQVSTRSQHTQP-TPEPSTAPSTSFLLPMGSPPPAEG 53
Db 2 LPQP-----PPHQQLLPQPPQP-----PPPPPPPPG 30

RESULT 7
Q9WTY7 PRELIMINARY; PRT; 54 AA.
AC Q9WTY7
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE DEUTHERIAL NITRIC OXIDE SYNTHASE (FRAGMENT).
GN NOS3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV129;
RX MEDLINE=99096466; PubMed=9878824;
RA Teichert A.M., Karantzoulis-Fegaras F., Wang Y., Mawji I.A., Bei X.,
RA Ganapandithen K., Marsden P.A.;
RT "Characterization of the murine endothelial nitric oxide synthase
RT promoter.";
RL Biochim. Biophys. Acta 1443:352-357(1998).
DR ENBL; AF091262; AAD22613.1; -.
FT NON_TER 54
SQ SEQUENCE 54 AA; 5193 MW; 1DA456A21958B2EA CRC64;

Query Match 16.7%; Score 50.5; DB 11; Length 54;
Best Local Similarity 47.8%; Pred. No. 96;
Matches 11; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 29 PTPPSTAPSTSFLLPMGSPSPA 51
Db 34 PAPEPSQAPA-----PPSPTRPA 51

RESULT 8
Q9GL38 PRELIMINARY; PRT; 52 AA.
AC Q9GL38
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE CALPASTATIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.

RA Chung H.Y., Davis M.E., Hines H.C.;
RT "PCR-SSCP analysis of the bovine calpastatin gene domain L region.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AY008267; AAG23869.1; -.
FT NON_TER 1
FT NON_TER 52
SQ SEQUENCE 52 AA; 5749 MW; CBD5A7449AFDDA89 CRC64;

Query Match 16.6%; Score 50; DB 6; Length 52;
Best Local Similarity 33.3%; Pred. No. 1e+02;
Matches 11; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

QY 5 PTRSMAPGAVHLPPQVSTRSQHTQPTPEPSTAP 37
Db 19 PKHSSDTGSKHAPKEKAVSKSSQPPSEKSTKP 51

RESULT 9
Q9J7C8 PRELIMINARY; PRT; 54 AA.
AC Q9J7C8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE LARGE T ANTIGEN (FRAGMENT).
OS Simian virus 40 (SV40).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC-028863B-1;
RA Rizzo P., Carbone M.;
RT "SV40 from 1955 commercial parental (Salk) poliovaccine.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF180738; AAF28272.1; -.
FT NON_TER 1
SQ SEQUENCE 54 AA; 5868 MW; 556CDAB682C1EFCD CRC64;

Query Match 16.4%; Score 49.5; DB 14; Length 54;
Best Local Similarity 32.7%; Pred. No. 1.2e+02;
Matches 16; Conservative 5; Mismatches 25; Indels 3; Gaps 2;

QY 8 SMAPGAVHLPPQ-VSTRSQHTQPTPEPSTAPSTSFLLPMGSPSPAEGST 55
Db 8 SQSQGSFQAPQPSQSQSXHDHNPVHICRGXTCXKKP--PTPPPEPET 54

RESULT 10
Q15218 PRELIMINARY; PRT; 46 AA.
AC Q15218
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE SALIVARY PROLINE-RICH PROTEIN 2 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298176; PubMed=6089212;
RA Azen E.A., Lyons K.M., McGonigal T., Barrett N.L., Clements L.S.,
RA Maeda N., Vanin E.F., Carlson D.M., Smithies O.;
RT "Clones from the human gene complex coding for salivary proline-rich
RT proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5561-5565(1984).
DR ENBL; K02578; AAA36505.1; -.
FT NON_TER 1
FT NON_TER 46
SQ SEQUENCE 46 AA; 4592 MW; FCE1D38D8DEDC173 CRC64;
```


Query Match 15.9%; Score 48; DB 4; Length 38;
Best Local Similarity 46.7%; Pred. No. 1.3e+02;
Matches 14; Conservative 2; Mismatches 6; Indels 8; Gaps 2;

QY 14 VHLPOPVSTRSQHTQTPPEPSTAPSTSFLL 43
| | | | |
DB 13 VEKQPQVSL-----PTPHN--PKSSQLL 34

RESULT 15

Q9GKJ3 PRELIMINARY; PRT; 42 AA.
AC Q9GKJ3;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE MYOSIN LIGHT CHAIN KINASE (FRAGMENT).
GN MYLK.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Van Poucke M., Yerle M., Tuggle C., Chardon P., Van Zeveren A.,
RA Peelman L.J.;
RT "Integration of porcine chromosome 13 maps."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF222917; AAG41130.1; -
KW Kinase.
FT NON_TER 1 1
FT NON_TER 42 42
SQ SEQUENCE 42 AA; 4677 MW; 78BDD867E66EF64F CRC64;

Query Match 15.9%; Score 48; DB 6; Length 42;
Best Local Similarity 36.6%; Pred. No. 1.4e+02;
Matches 15; Conservative 2; Mismatches 14; Indels 10; Gaps 2;

QY 14 VHLPOPVSTRS---OHTQTPPEPSTAPSTSFLLPMGSPSP 50
| | | | | : | | | | |
DB 8 VHSPOQVDFRSVLARKGTPKTPVEKLP-----PPKPPTP 42

Search completed: July 13, 2001, 17:43:47
Job time: 164 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:38:03 ; Search time 19.91 Seconds
(without alignments)
173.559 Million cell updates/sec

Title: US-09-800-909-2_COPY_201_257

Perfect score: 302

Sequence: 1 TSTSPTRSMAPGAVHLPQPV.....STSELLPMGPSPRAEGSTGD 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 222704

Minimum DB seq length: 0

Maximum DB seq length: 57

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0601.*

```
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	22.7	47	21 AAB56300	Human secreted pro
2	67.5	22.4	49	19 AAW59911	Amino acid sequenc
3	59	19.5	45	12 AAR15261	Linking B region #
4	58	19.2	54	15 AAR49722	Sequence of a pept
5	58	19.2	54	15 AAR49536	Camel Ig 2-heavy c
6	56.5	18.7	41	20 AAY01285	Peptide encoded by
7	55	18.2	46	20 AAW8522	Amphotropic hyperv
8	55	18.2	52	22 AAB61193	Human INTERCEPT 21
9	54.5	18.0	55	21 AAB42690	Human OREF ORF2454
10	54	17.9	21	9 AAP81606	Sequence of human
11	54	17.9	36	21 AAY65289	Human 5' EST relat

12	53.5	17.7	37	18	AAW26641	H. insolens family	
13	52.5	17.4	21	14	AAR37992	Heavy chain hinge	
14	52	17.2	50	18	AAW35739	Mucin peptide MUC1	
15	52	17.2	51	18	AAR31697	Mucin peptide MUC1	
16	51	16.9	40	17	AAR88860	Peptide targeted	
17	51	16.9	43	20	AAY49096	Amino acid sequenc	
18	50	16.6	16	17	AAR99693	IgA protease speci	
19	50	16.6	19	20	AAW99681	Human IgA1 hinge r	
20	50	16.6	21	20	AAY42923	Core peptide in th	
21	50	16.6	25	19	AAW64623	Human IgA1 CH2 reg	
22	50	16.6	25	19	AAW61595	Human IgA1 CH2 rec	
23	50	16.6	25	20	AAW85774	IgA1 CH2 region pr	
24	50	16.6	52	21	AAG03497	Human secreted pro	
25	49.5	16.4	40	16	AAR68002	Mucin repeat sequ	
26	49.5	16.4	40	19	AAW72703	Human mucin peptid	
27	49.5	16.4	41	12	AAR15255	Linking B region #	
28	49	16.2	27	9	AAP80541	Sequence of portio	
29	49	16.2	35	15	AAW49650	Sequence of hinge	
30	49	16.2	35	15	AAW49529	Camel 2-chain Ig h	
31	49	16.2	40	19	AAW54873	Carcinoma-associat	
32	49	16.2	46	20	AAY01284	Peptide encoded by	
33	49	16.2	47	21	AAB38081	Fragment of human	
34	48.5	16.1	34	20	AAY03809	Lck binding active	
35	48.5	16.1	45	21	AAG35605	Arabidopsis thalia	
36	48.5	16.1	53	21	AAG00350	Human secreted pro	
37	48.5	16.1	57	20	AAY12639	Human 5' EST seque	
38	48	15.9	39	20	AAW74023	Human sucrose isom	
39	-	48	15.9	41	2	AAP10196	Sequence of struct
40	48	15.9	41	20	AAW74094	Gastro-intestinal	
41	48	15.9	41	21	AAW96170	MUC1 repeat sequen	
42	48	15.9	42	8	AAP71006	Sequence of fragme	
43	48	15.9	42	9	AAP80040	Beta-human chorion	
44	48	15.9	42	10	AAP91846	Analogue of beta s	
45	48	15.9	42	19	AAW69459	HCG antigenic pept	

ALIGNMENTS

```
RESULT 1
AAB56300
ID AAB56300 standard; Protein; 47 AA.
XX
AC AAB56300;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 99 SEQ ID NO:394.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antithratic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; pathological condition;
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW neoplasia; cardiovascular disorder; cardiac arrest; cerebral ischaemia;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; skin aging; food additive; preservative.
XX
OS Homo sapiens.
XX
PN WO200070042-A1.
XX
PD 23-NOV-2000.
XX
PF 11-MAY-2000; 2000WO-US12788.
XX
PR 13-MAY-1999; 99US-0134068.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;
PI Duan RD, Florence KA, Soppet DR;
```

```

XX WPI; 2000-679828/66.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition -
XX
XX Disclosure; Page 1035; 1065pp; English.
XX
XX The polynucleotide sequences given in AAC99818 to AAC99977 encode the
XX human secreted proteins given in AAB56077 to AAB56362. Human secreted
XX proteins have activities based on the tissues and cells the genes are
XX expressed in. Examples of activities include: immunosuppressive;
XX antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
XX vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
XX virucide; fungicide; and ophthalmological. The human secreted
XX polynucleotides and proteins can be used to prevent, treat or ameliorate
XX a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
XX dogs, chickens or sheep. They are also used in diagnosing a pathological
XX condition or susceptibility to a pathological condition. Disorders which
XX are diagnosed or treated include autoimmune diseases e.g. rheumatoid
XX arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
XX liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
XX disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
XX e.g. Alzheimer's disease, infections caused by bacteria, viruses and
XX fungi and ocular disorders e.g. corneal infection. The proteins can also
XX be used to aid wound healing and epithelial cell proliferation, to
XX prevent skin aging due to sunburn, to maintain organs before
XX transplantation, for supporting cell culture of primary tissues, to
XX regenerate tissues and in chemotaxis. The proteins can also be used as a
XX food additive or preservative to increase or decrease storage
XX capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used
XX in the exemplification of the present invention.
XX
XX Sequence 47 AA;
XX
XX Query Match 22.7%; Score 68.5; DB 21; Length 47;
XX Best Local Similarity 40.5%; Pred. NO. 2.4;
XX Matches 17; Conservative 3; Mismatches 17; Indels 5; Gaps 1
XX
XX QY 19 PVSRSQHTQTPPE-----PSTAPSTSELLPMGSPPPARGST 55
XX |||| ||| | : : | | | |||| | :
XX 5 pvsitlpgsptshahbptcpspbtpxshpsxpptatss 46
XX
XX RESULT 2
XX AAW59911
XX ID AAW59911 standard; peptide: 49 AA.
XX AC AAW59911;
XX AD
XX AE
XX AF
XX AG 20-NOV-1998 (first entry)
XX AH
XX AI Amino acid sequence of the mutanase enzyme PT box.
XX AJ
XX AK Mutanase enzyme; PT box; alpha-1,3 glucoside bond; mutan; plaque;
XX AL bacteria; teeth.
XX AM
XX AN Bacillus sp.
XX AO
XX AP
XX AQ JF10201483-A.
XX AR
XX AS 04-AUG-1998.
XX AT
XX AU 01-OCT-1997; 97JP-0284362.
XX AV
XX AW 25-NOV-1996; 96JP-0314057.
XX AX
XX AY (LLOY ) LION CORP.
XX AZ
XX BA 1998-474495/41.
XX BB
XX BC Gene encoding a mutanase enzyme - used for prevention and removal
XX of plaque.

```


XX SQ Sequence 52 AA;

Query Match 18.2%; Score 55; DB 22; Length 52;
 Best Local Similarity 35.9%; Pred. No. 56;
 Matches 14; Conservative 5; Mismatches 18; Indels 2; Gaps 1;

QY 19 PVSTRSQHTQTP--EPSTAPSTFLLPMGPPPPAGST 55
 I : : ||| | | : | : | | | : ||
 Db 12 pllipatqtppraephkssvlsttpdpapsqgqast 50

RESULT 9
 AAB42690
 ID AAB42690 standard; Protein; 55 AA.
 XX AC AAB42690;
 DT 08-FEB-2001 (first entry)
 XX DE Human ORFX ORF2454 polypeptide sequence SEQ ID NO:4908.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antirheumatic; antithyroid;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.
 OS
 XX WO200058473-A2.
 PN
 PD 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 XX
 XX 31-MAR-1999; 99US-0127607.
 PR
 PR 02-APR-1999; 99US-0127636.
 PR
 PR 05-APR-1999; 99US-0127728.
 PR
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 XX Shinkets RA, Leach M;
 PI
 XX WPI; 2000-602362/57.
 DR
 DR N-PSDB; AAC76899.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 4083; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining

CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 55 AA;

Query Match 18.0%; Score 54.5; DB 21; Length 55;
 Best Local Similarity 38.3%; Pred. No. 67;
 Matches 18; Conservative 1; Mismatches 17; Indels 11; Gaps 3;

QY 19 PVSTRSQHTQTPPEPS-----TAPSTFLLPMGPS---PPAEGSTG 56
 I : : ||| | | : | | : | | : ||| | |
 Db 5 psgrtprtptpycpghdrllppsrp--lpagpasafppaersrg 49

RESULT 10
 AAP81606
 ID AAP81606 standard; Protein; 21 AA.
 XX AC AAP81606;
 XX 02-OCT-1990 (first entry)
 DT
 XX Sequence of human immunoglobulin A1 (IgA1).
 DE
 XX Human Cu/Zn superoxide dismutase polymer; PSODCF1SODHAL;
 KW human immunoglobulin A1; hinge region; antinflammatory agent;
 KW
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Region 1..20
 FT Protein /note="Hinge #1"
 FT 4..20
 FT /note="IgA1"
 XX
 XX EP283244-A.
 PN
 XX 21-SEP-1988.
 PD
 XX 15-MAR-1988; 88EP-0302244.
 PF
 XX 16-MAR-1987; 87US-0026143.
 PR
 XX (CHIR-) CHIRON CORP.
 PA
 XX Hallelwell RA, Mullenbach G;
 PI
 XX WPI; 1988-265657/38.
 DR
 DR P-PSDB; AAN80435.
 XX
 XX Superoxide dismutase polymers having extended in vivo life -
 PT comprising superoxide dismutase monomers covalently coupled by
 PT polypeptide spacers
 PT
 XX Example; Page 5; 2lpp; English.

XX A Cu/Zn superoxide dismutase (SOD) polymer of SOD units of at least 2
 CC SOD monomers covalently coupled, carboxy terminus to amino terminus, to
 CC each other by a polypeptide spacer of at least 3 amino acids is claimed.
 CC The polymers are of formula
 CC (SOD monomer - IgA - SOD monomer)x or
 CC (SOD monomer - IgA - SOD monomer - IgA - SOD monomer)x


```
XX 20-MAR-1997; 97WO-US04493.
XX
XX 20-MAR-1996; 96US-0013775.
XX
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
XX Livingston PO, Zhang S;
XX
XX WPI; 1997-480158/44.
XX
XX Vaccine effective against cancer of the breast, prostate, colon,
XX lung or pancreas - comprising mucin peptide, especially MUC1,
XX conjugated to immunogenic protein, especially keyhole limpet
XX haemocyanin
XX
XX Disclosure; Page 11; 45pp; English.
XX
XX This mucin peptide is used in a vaccine capable of producing an immune
XX response which recognises a mucin. The mucin peptide is selected from
XX MUC1 peptide group. The vaccine comprises an amount of the mucin peptide
XX conjugated to an immunogenic protein effective to stimulate or enhance
XX immune response in the subject, together with an adjuvant and a vehicle.
XX A cysteine is added to the N-terminal of this peptide to facilitate the
XX conjugation with protein carriers. The immunogenic protein is a keyhole
XX limpet haemocyanin (KLH) or its derivative. The vaccine can be used to
XX induce an immune response in patients suffering from a cancer of the
XX type where the cancer cells have mucin on their surface, e.g. breast
XX cancer, prostate cancer, lung cancer, colon cancer or pancreas cancer.
XX
XX Sequence 51 AA;
```

Query Match 17.2%; Score 52; DB 18; Length 51;
Best Local Similarity 36.4%; Pred. No. 1.le+02;
Matches 20; Conservative 3; Mismatches 18; Indels 14; Gaps 3;

QY 1 TSTSPTRSMAPGAVHLPQVPVSTRSQHTQPTPEPSTAPSTSFLLPMGPPPAEGST 55
|| || |||: | : | : | ||| |
Db 3 tsapdtrp-aggstappahgvtapdtrp-pgsta-----ppahgvt 43

Search completed: July 13, 2001, 17:40:58
Job time: 175 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:38:48 ; Search time 12.11 Seconds
(without alignments)
94.818 Million cell updates/sec

Title: US-09-800-909-2_COPY_201_257

Perfect score: 302

Sequence: 1 TTSPTSRMAPGAVHLPQPV.....STSFLLPMGPPPAEGSTGD 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 139220

Minimum DB seq length: 0

Maximum DB seq length: 57

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	27.5	15	1	US-08-221-583-56
2	83	27.5	15	5	PCT-US95-04018-56
3	82	27.2	15	1	US-08-221-583-58
4	82	27.2	15	5	PCT-US95-04018-58
5	81	26.8	15	1	US-08-221-583-57
6	81	26.8	15	1	US-08-221-583-62
7	81	26.8	15	5	PCT-US95-04018-57
8	81	26.8	15	5	PCT-US95-04018-62
9	80	26.5	15	1	US-08-221-583-55
10	80	26.5	15	1	US-08-221-583-59
11	80	26.5	15	5	PCT-US95-04018-59
12	80	26.5	15	5	PCT-US95-04018-55
13	79	26.2	15	1	US-08-221-583-61
14	79	26.2	15	5	PCT-US95-04018-61
15	78	25.8	15	1	US-08-221-583-60
16	78	25.8	15	5	PCT-US95-04018-60
17	69	22.8	15	1	US-08-221-583-54
18	69	22.8	15	5	PCT-US95-04018-54
19	59	19.5	45	1	US-08-361-920-19
20	59	19.5	45	1	US-08-479-939-19
21	59	19.5	45	1	US-08-483-432-19
22	58	19.2	54	1	US-08-471-780C-44
23	58	19.2	54	1	US-08-467-282B-44
24	58	19.2	54	2	US-08-471-282A-44
25	58	19.2	54	2	US-08-466-710C-44
26	58	19.2	54	3	US-08-468-739C-44
27	55	18.2	15	1	US-08-221-583-53

28 55 18.2 15 5 PCT-US95-04018-53 Sequence 53, Appl
29 55 18.2 46 3 US-08-856-074A-39 Sequence 39, Appl
30 55 18.2 48 6 5171685-7 Patent No. 5171685
31 55 18.2 48 6 5518916-7 Patent No. 5518916
32 53.5 17.7 37 3 US-08-814-052-37 Sequence 37, Appl
33 53.5 17.7 37 3 US-08-812-829-29 Sequence 29, Appl
34 52.5 17.4 21 5 PCT-US92-10432-1 Sequence 1, Appl
35 52.5 17.4 22 1 US-08-442-542-43 Sequence 43, Appl
36 52.5 17.4 22 3 US-08-765-469-43 Sequence 43, Appl
37 50.5 16.7 53 6 5422248-4 Patent No. 5422248
38 50 16.6 16 1 US-08-366-591-15 Sequence 15, Appl
39 49.5 16.4 33 1 US-08-237-716-11 Sequence 11, Appl
40 49.5 16.4 40 1 US-08-099-354-1 Sequence 1, Appl
41 49.5 16.4 40 2 US-08-288-059-7 Sequence 7, Appl
42 49.5 16.4 41 1 US-08-361-920-13 Sequence 13, Appl
43 49.5 16.4 41 1 US-08-479-939-13 Sequence 13, Appl
44 49.5 16.4 41 1 US-08-483-432-13 Sequence 13, Appl
45 49 16.2 16 1 US-08-471-033-47 Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-08-221-583-56
; Sequence 56, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-56

Query Match 27.5%; Score 83; DB 1; Length 15;
Best Local Similarity 100.0%; Pred.No. 0.0034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 HLPQPVSTRSQHTQP 29
Db 1 HLPQPVSTRSQHTQP 15

RESULT 2

```
PCT-US95-04018-56
; Sequence 56, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-04018-56

Query Match 27.5%; Score 83; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 HLPQPVSTRSQHTQP 29
Db 1 HLPQPVSTRSQHTQP 15
|||||

RESULT 3
US-08-221-583-58
; Sequence 58, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595Fris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania

PCT-US95-04018-56
; Sequence 56, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-04018-56

Query Match 27.2%; Score 82; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 SQHTQPTPEPSTAPS 38
Db 1 SQHTQPTPEPSTAPS 15
|||||

RESULT 4
PCT-US95-04018-58
; Sequence 58, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
```

NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-58

Query Match 27.2%; Score 82; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 SQHTQPTPEPSTAPS 38
Db 1 SQHTQPTPEPSTAPS 15

RESULT 5
US-08-221-583-57
Sequence 57, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-57

Query Match 26.8%; Score 81; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 STRSQHTQPTPEPST 35
Db 1 STRSQHTQPTPEPST 15

RESULT 6
US-08-221-583-62
Sequence 62, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-62

Query Match 26.8%; Score 81; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 APSTSFLLPMGPSP 50
Db 1 APSTSFLLPMGPSP 15

RESULT 7
PCT-US95-04018-57
Sequence 57, Application PC/TUS9504018
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
APPLICANT: Kruszynski, Marian
APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: Norris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04018
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: US 08/221,580
APPLICATION NUMBER: US 08/221,580
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA: US 08/221,583
APPLICATION NUMBER: US 08/221,583
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA: US 08/221,581
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-57

Query Match 26.8%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 STRSQHTQTPPEST 35
Db 1 STRSQHTQTPPEST 15

RESULT 8
US-08-221-583-62
Sequence 62, Application PC/TUS9504018
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
APPLICANT: Kruszyński, Marian
APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04018
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,580
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA: US 08/221,583
APPLICATION NUMBER: US 08/221,583
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA: US 08/221,581
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-62

Query Match 26.8%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 APSTSFLLPMGPSP 50
Db 1 APSTSFLLPMGPSP 15

RESULT 9
US-08-221-583-55
Sequence 55, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25.mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-55

Query Match 26.5%; Score 80; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 GAVHLPQVSTRSQH 26
Db 1 GAVHLPQVSTRSQH 15

RESULT 10
US-08-221-583-59
; Sequence 59, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25:mdtcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221.583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: COOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-59

Query Match 26.5%; Score 80; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 TQTPPEPSTAPSTSF 41
|||||
DB 1 TQTPPEPSTAPSTSF 15

RESULT 11
PCT-US95-04018-55
; Sequence 55, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-55

Query Match 26.5%; Score 80; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 GAVHLPQPVSTRSQH 26
|||||
DB 1 GAVHLPQPVSTRSQH 15

RESULT 12
PCT-US95-04018-59
; Sequence 59, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581

FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-59

Query Match 26.5%; Score 80; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 TQTPPEPSTAPSTF 41
Db 1 TQTPPEPSTAPSTF 15

RESULT 13
US-08-221-583-61
Sequence 61, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-61

Query Match 26.2%; Score 79; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 PSTAPSTSFLLPMGP 47
Db 1 PSTAPSTSFLLPMGP 15

RESULT 14
PCT-US95-04018-61
Sequence 61, Application PC/TUS9504018
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
APPLICANT: Kruszynski, Marian
APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04018
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,580
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,583
FILING DATE: 01-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,581
FILING DATE: 01-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-61

Query Match 26.2%; Score 79; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 PSTAPSTSFLLPMGP 47
Db 1 PSTAPSTSFLLPMGP 15

RESULT 15
US-08-221-583-60
Sequence 60, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris

STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25.mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: COOR-0185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-60

Query Match 25.88; Score 78; DB 1; Length 15;
Best Local Similarity 100.00; Pred. No. 0.011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 30 TPEPSTAPSTFLLP 44
Db 1 TPEPSTAPSTFLLP 15

Search completed: July 13, 2001, 17:41:37
Job time: 169 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	71	23.9		54	2	A60726	mucin, gallbladder
2	56.5	19.0		52	2	S63324	hypothetical prote
3	55.5	18.7		42	2	I70082	glycoprotein Ib al
4	51	17.2		25	2	I40692	cena protein (Igale
5	50	16.8		46	2	S07073	arabinogalactan pr
6	47	15.8		31	2	T36022	small hypothetical
7	44	14.8		28	2	I56139	MHC class I HLA-J
8	43.5	14.6		47	2	S32108	sepiapterin reduct
9	43.5	14.6		34	2	D81737	hypothetical prote
10	43	14.5		51	2	I54315	pre-B cell Ig lamb
11	43	14.5		39	2	A46662	collagen alpha 2(V
12	42	14.1		36	2	A37172	collagen alpha 1(X
13	42	14.1		48	2	I46522	troponin T 2fa - r
14	42	14.1		55	2	S08424	H+-transporting AT
15	42	14.1		55	2	T11184	H+-transporting AT
16	41	13.8		42	2	T07030	extensin - tomato
17	40.5	13.6		50	2	A29789	mucin - sheep (fra
18	40	13.5		29	2	G39690	neural cell adhesi
19	40	13.5		50	2	H64801	hypothetical prote
20	40	13.5		51	2	A33756	dorsal protein - f
21	40	13.5		51	2	G72801	gp18 protein - Myc
22	40	13.5		52	2	S58216	hypothetical prote
23	40	13.5		54	2	T11131	ATP synthase subun
24	40	13.5		55	2	T11105	H+-transporting AT
25	39.5	13.3		52	2	S01945	myosin catalytic l
26	39	13.1		27	2	S11176	aspartate transcar
27	38.5	13.0		51	2	C81117	hypothetical prote
28	38	12.8		53	2	S23202	kappa-casein - bov
29	38	12.8		54	2	S14338	serpin I - horse (

QY 7 SMAPGAVHLPPQPVSTRSQHTQTPPEPSTAPSTSFLPMGPSP 48
 || :||| | |||| | | :|
 Db 11 SMOYSDIYIPTPPTHTHT-PTPHPH--PHTHTHTHNNNP 49

C; Keywords: hydroxyproline
F:6,7,9,11,14,16,18,20,26,30,34,36,39/Modified site: hydroxyproline (Pro) #status exp

Query Match 16.8%; Score 50; DB 2; Length 46;
Best Local Similarity 37.1%; Pred. No. 1.3e+02;
Matches 13; Conservative 4; Mismatches 16; Indels

QY 16 PQPVSTRSQHTQTPPEPSTAPSTSFLLPMGPPSPA 50
| : : | | | | : | | | |
Db 6 PPAPAPKAPAPAPVPEASTAPVAA--PTTXPSPPA 38

RESULT 6

T36022
small hypothetical protein SCC54.09c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change
C:Accession: T36022
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21581
A:Accession: T36022
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1..31 <SE>
A:Cross-references: EMBL:AL035591; PIDN:CAB38139.1; GSPDB:GN00
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCC54.09c

Query Match 15.8%; Score 47; DB 2; Length 31;
Best Local Similarity 30.4%; Pred. No. 1.6e+02;
Matches 7; Conservative 6; Mismatches 10; Indels

```
QY      18 PVSTRSQHTQTPPEPSTAPSTSF 40  
         |::|:::||::||::|  
Db       2 PLAARNEDNEPVPAPTVVVTGTGF 24
```

```

RESULT      7
I56139      MHC class I HLA-J antigen - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I56139; I72807
J:Mesner, G.; Zemmour, J.; Orr, H.T.; Parham, P.; Weiss, E.H.; Girdlestone, J.
R. Immunol. 148, 4043-4053, 1992
A:Title: HLA-J, a second inactivated class I HLA gene related to HLA-G and HLA-A. Imp
A:Reference number: I56139; MUID:92291530

```

A: Accession: J30139
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-28 <RES>
A: Cross-references: GB:M80468; NID:g188481; PIDN:AAA36306.1; PID:g188482
A: Accession: I72807
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-28 <RES>
A: Cross-references: GB:M80469; NID:g188483; PIDN:AAA36307.1; PID:g188484
C: Genetics:
C: Introns: 24/3
C: Superfamily: class I histocompatibility antigen; immunoglobulin homology

```
Query Match      14.8%   Score 44; DB 2; Length 28;
Best Local Similarity 40.9%; Pred. NO. 2.6e+02;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
```

RESULT 3
I70082
glycoprotein Ib alpha variant B - human (fragment)
C;Species: Homo sapiens. (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I70082
R;Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.
J. Biol. Chem. 267, 10055-10061, 1992
A;title: Polymorphism of human glycoprotein Ib alpha results from a variable number of t
ations.

A; Accession: I70082
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-42 <RES>
A; Cross-references: GB:S34439; NID:g249178; PID: AAB22153.1; PID: g249179
C; Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein
C; Keywords: glycoprotein

Query Match	18.7%	Score	55.5;	DB 2:	Length	42;			
Best Local Similarity	35.6%	Pred. No.	37;						
Matches	16;	Conservative	6;	Mismatches	18;	Indels	5;	Gaps	3;

QY 4 PTRMAPGAVHLPPQPVSTRSQHTQPTPEPSTAPSTSFLPLMGPPS 48
 || || : || : ||||: ||: ||: ||
 Db 2 PTSEPAAPSPT-TPEPTSEPA-SPTTPEPTSEPA-S--PTTPEP 41

RESULT 4
I40692
cenaA protein (IqAlh) - Cellulomonas fimi (fragment)
C:Species: Cellulomonas fimi
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 22-Oct-1999
C:Accession: I40692
R:Miller, P.B.; Shen, H.; Gilkes, N.R.; Kilburn, D.G.; Miller, R.C.; Plaut, A.G.; Warren
FEMS Microbiol. Lett. 92, 199-204, 1992
A:Title: Endoglucanase A from Cellulomonas fimi in which the hinge sequence of human IgA
hoseae.
A:Reference number: I40692
A:Accession: I40692
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25 <RES>
A:Cross-references: EMBL:X65780; NID:g312035; PIDN:CAA46663.1; PID:g312036

Query Match	17.2%	Score 51;	DB 2;	Length 25;
Best Local Similarity	46.2%	Pred. No. 55;		
Matches 12;	Conservative	1;	Mismatches	8;
Matches 5;	Indels	5;	Gaps	1;

QY 23 SQHTQPTPEPSTAPSTSFLLPMGPSP 48
↑ | | | | | : | | |
Db 2 SVSTPPTPSPSTPPT-----PSP 19

RESULT 5
S07073
arabinogalactan protein - Italian ryegrass (fragments)
C:Species: Lolium multiflorum (Italian ryegrass)
C:Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 05-Dec-1998
C:Accession: S07073
R:Gleeson, P.A.; McNamara, M.; Wettenhall, R.E.H.; Stone, B.A.; Fincher, G.B.
Biochem. J. 264, 857-862, 1989
A:Title: Characterisation of the hydroxyproline-rich protein core of an arabinogalactan-
A:Reference number: S07073; MUID:90147544
A:Accession: S07073
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 4-46 <GLE>
A>Note: 19-His and 23-Leu were also found

```

RESULT 8
S32108
sepiapterin reductase (EC 1.1.1.153) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 12-Sep-1997
C:Accession: S32108
R:Maier, J.
submitted to the EMBL Data Library, March 1993
A:Reference number: S32108
A:Accession: S32108
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-47 <MAI>
A:Cross-references: EMBL:Z21947
C:Keywords: oxidoreductase

Query Match 14.6%; Score 43.5; DB 2; Length 47;
Best Local Similarity 39.3%; Pred. No. 4.9e+02;
Matches 11; Conservative 3; Mismatches 13; Indels 1; Gaps 1;

QY 19 VSTRQHTOPTPEPTAPSTFLLPMGP 46
|| | | | | | | | | | | | | | |
Db 14 VSGLKSHSPAPHP-TSTSAVCSLTIGP 40

RESULT 9
D81737
hypothetical protein TC0127 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: D81737
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwin, M.C.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255
A:Accession: D81737
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-34 <TET>
A:Cross-references: GB:AE002280; GB:AE002160; NID:g7190162; PIDN:AAF39005.1; PID:g719016
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0127

Query Match 14.6%; Score 43.5; DB 2; Length 54;
Best Local Similarity 42.9%; Pred. No. 5.7e+02;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 27 QTPPTAPSTFLLPMGPSPPAEGST 54
|| | | | | | | | | | | | | | |
Db 4 QPSKPSQASSLS-LKGGDPAPPSPGRAT 30

RESULT 10
154515
pre-B cell Ig lambda-like omega light chain (non-rearranging) IGLL2 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 23-Jul-1999
C:Accession: I54515
R:Bauer, T.R.; McDermid, H.E.; Budarf, M.L.; Van Keuren, M.L.; Blomberg, B.B.
Immunogenetics 38, 387-399, 1993
A:Title: Physical location of the human immunoglobulin lambda-like genes, 14.1, 16.1, and
A:Reference number: I54515; MUID:94011089
A:Accession: I54515
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-31 <RES>
A:Cross-references: GB:L02326; NID:g292400; PIDN:AAA16174.1; PID:g451281

C:Genetics:
A:Gene: GDB:IGLL2
A:Cross-references: GDB:137325
A:Map position: 22q11.2-22q11.2
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 14.5%; Score 43; DB 2; Length 31;
Best Local Similarity 37.5%; Pred. No. 3.5e+02;
Matches 9; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 31 EPSTAPSTFLLPMGPSPPAEGST 54
|| | | | | | | | | | | | | | |
Db 2 QPKTTPSVILFLPSCPEPQANKAT 25

RESULT 11
A46662
collagen alpha 2(V) chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Jul-1995
C:Accession: A46662
R:Mayne, R.; Brewton, R.G.; Mayne, P.M.; Baker, J.R.
J. Biol. Chem. 268, 9381-9386, 1993
A:Title: Isolation and characterization of the chains of type V/type XI collagen pres
A:Reference number: A46662; MUID:93252802
A:Accession: A46662
A:Molecule type: protein
A:Residues: 1-39 <MAY>
A:Experimental source: vitreous humor
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

Query Match 14.5%; Score 43; DB 2; Length 39;
Best Local Similarity 53.3%; Pred. No. 4.3e+02;
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 42 LPMGPSPPAEGSTGD 56
|| | | | | | | | | | | | | | |
Db 14 LPGPDPGPPGEAGPGD 28

RESULT 12
A37172
collagen alpha 1(XII) chain-like, skin and tendon - bovine (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 19-Oct-1995
C:Accession: A37172
R:Dublet, B.; Van Der Rest, M.
Ann. N.Y. Acad. Sci. 580, 436-439, 1989
A:Title: Comparison between chicken type XII collagen and bovine homologues.
A:Reference number: A37172
A:Accession: A37172
A:Molecule type: protein
A:Residues: 1-36 <DUB>

Query Match 14.1%; Score 42; DB 2; Length 36;
Best Local Similarity 41.4%; Pred. No. 5.1e+02;
Matches 12; Conservative 3; Mismatches 8; Indels 6; Gaps 3;

QY 24 QHTQTPPEPSTA-PSTSFLLPMGPSPPAE 51
|| | | | | | | | | | | | | | |
Db 3 QERSP-PQANAVPSX----PASPSPLQ 26

RESULT 13
146522
troponin T 2fa - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
```

C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999

C;Accession: I46522
J;Briggs, M.M.; Lin, J.J.; Schachat, F.H.
R. Muscle Res. Cell. Motil. 8, 1-12, 1987
A;Title: The extent of amino-terminal heterogeneity in rabbit fast skeletal muscle troponin T
A;Reference number: I46522; MUID:87251333
A;Accession: I46522
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-48 <BRI>
A;Cross-references: EMBL:U04975; NID:g440810; PIDN:AAAL6028.1; PID:g440811
C;Superfamily: troponin T

Query Match 14.1%; Score 42; DB 2; Length 48;
Best Local Similarity 37.9%; Pred. No. 6.8e+02;
Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 2;

QY 9 APGAVHLPPQVPVSTRSQHTQPTPEPS-TAP 36
: | | | | : | | | |
Db 25 SPAEVHEPEEV-----HEEEKPRPKLTAP 48

RESULT 14

S08424
H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - Atlantic cod mitochondrion
C:Species: mitochondrion Gadus morhua (Atlantic cod)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 20-Jun-2000
C:Accession: S08424; T11824
R:Johansen, S.; Guddal, P.H.; Johansen, T.
Nucleic Acids Res. 18, 411-419, 1990
A:Title: Organization of the mitochondrial genome of Atlantic cod, Gadus morhua.
A:Reference number: S08424; MUID:90174958

A;Accession: S08424
A;Molecule type: DNA
A;Residues: 1-55 <JGH>
A;Cross-references: EMBL:X17659; NID:g14006; PIDN:CAA35655.1; PID:g14007
R;Johansen, S.; Bakke, I.
Mol. Marine Biol. Biotechnol. 5, 203-214, 1996
A;Title: The complete mitochondrial DNA sequence of Atlantic cod, *Gadus morhua*: Relevant
A;Reference number: Z17351; MUID:96414925

A;Accession: I11024
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-55 <J02>
A;Cross-references: EMBL:X99772; PIDN:CAA68110.1
A;Experimental source: norwegian coastal stock (NC-1)
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCL
C;Superfamily: H+-transporting ATP synthase protein 8
C;Keywords: ATP biosynthesis; hvdrolase; membrane-associated complex; mitochondrion; oxygen

```
Query Match      14.1%; Score 42; DB 2; Length 55;
Best Local Similarity 32.0%; Pred. NO. 7.9e+02;
Matches 8; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
```

QY 15 LPQPVSTRSQHTQTPPEPSTAPSTS 39
|| | : : : : :
Db 24 LPPKVMHTFPNEPSPQGMTPKTA 48

RESULT 15

Tll1184
H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - Smithornis sharpei mitochondrion
C:Species: mitochondrion Smithornis sharpei
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 07-Dec-1999
C:Accession: Tll1184
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: Z17242

A: Accession: T11184
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-55 <MIN>
A: Cross-references: EMBL:AF090340; NID:g4894488; PID:g4894493; PIDN:AAD32518.1
C: Genetics:
A: Genome: mitochondrion
A: Genetic code: SGC1
C: Superfamily: H+-transporting ATP synthase protein 8
C: Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion;

```

Query Match      14.1%;
Best Local Similarity 52.4%;
Matches 11; Conservative 1

QY 28 PTPEPSTAPSTSFLLPMGPSP 48
Db 33 PTNPPSTKTST-----MKPSP 49

```

Search completed: July 13, 2001, 17:45:09
Job time: 106 sec

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Mitochondrion.


```
CC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadoidei; Gadidae;
CC Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NORWEGIAN COASTAL 1; TISSUE=Liver;
RX MEDLINE=90174958; PubMed=2308841;
RA Johansen S., Guddal P.H., Johansen T.;
RT "Organization of the mitochondrial genome of Atlantic cod, Gadus
RT morhua.";
RL Nucleic Acids Res. 18:411-419(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NORWEGIAN COASTAL 1;
RX MEDLINE=96414925; PubMed=8817926;
RA Johansen S., Bakke I.;
RT "The complete mitochondrial DNA sequence of Atlantic cod (Gadus
RT morhua): relevance to taxonomic studies among codfishes.";
RL Mol. Mar. Biol. Biotechnol. 5:203-214(1996).
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17659; CAA35655.1; -
CC EMBL; X93772; -; NOT_ANNOTATED_CDS.
DR PIR; S08424;
DR InterPro; IPR001421; -
DR Pfam; PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 4 24 POTENTIAL.
SQ SEQUENCE 55 AA; 6481 MW; E85C81E63DB48B15 CRC64;
-----
Query Match 14.1%; Score 42; DB 1; Length 55;
Best Local Similarity 32.0%; Pred. No. 5e+02;
Matches 8; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
-----
QY 15 LPQPVSTRSQHTQTPPEPSTAPSTS 39
| | : : : : :
Db 24 LPPKVMATFPNEPSPQGMTPKTA 48
-----
RESULT 3
ID VG18_BPMD2 STANDARD; PRT; 51 AA.
AC O64211;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE GENE 18 PROTEIN (GP18).
GN 18.
OS Mycobacteriophage D29.
OC Viruses.
OX NCBI_TaxID=28369;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: Implications for phage
RT evolution.";
RL J. Mol. Biol. 279:143-164(1998).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF022214; AAC18450.1; -
CC EMBL; AF022214; AAC18450.1; -
SQ SEQUENCE 51 AA; 5680 MW; EF85B1AFF5786A34 CRC64;
-----
Query Match 13.5%; Score 40; DB 1; Length 51;
Best Local Similarity 33.3%; Pred. No. 6.9e+02;
Matches 11; Conservative 4; Mismatches 14; Indels 4; Gaps 1;
-----
QY 6 RSMAPGAVHLPPQVSTRSQHTQTPPEPSTAPST 38
| | : : : : :
Db 21 RLIIAGGWHKRPKRPRT---TKPKAPKQEPAT 49
-----
RESULT 4
ID ATP8_PELSU STANDARD; PRT; 55 AA.
AC O79674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
GN MTATP8 OR ATP8.
OS Pelomedusa subrufa (African side-necked turtle).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Pleurodira; Pelomedusidae; Pelomedusa.
OX NCBI_TaxID=44522;
RN [1]
RP SEQUENCE FROM N.A.
RA zardoya R.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF039066; AAD05054.1; -
CC EMBL; AF039066; AAD05054.1; -
DR InterPro; IPR001421; -
DR Pfam; PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 4 24 POTENTIAL.
SQ SEQUENCE 55 AA; 6536 MW; D8D4EC8F8651A001 CRC64;
-----
Query Match 13.5%; Score 40; DB 1; Length 55;
Best Local Similarity 36.0%; Pred. No. 7.4e+02;
Matches 9; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
-----
QY 14 HLPQVSTRSQHTQTPPEPSTAPST 38
| | : : : : :
Db 31 HIPNNSPTNKKNMLTTPMPWTWPT 55
-----
RESULT 5
ID TAGP_HUMAN STANDARD; PRT; 56 AA.
AC Q9Y3F1;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
```


Query Match	12.3%;	Score 36.5;	DB 1;	Length 55;
Best Local Similarity	40.5%;	Pred. NO. 1.5e+03;		
Matches 15;	Conservative 2;	Mismatches 17;	Indels 3;	Gaps 2;

RP SEQUENCE.
 RX MEDLINE=70252720; pubmed=4194291;
 RA Bornstein P.; Nesse R.;

RT The comparative biochemistry of collagen: the structure of rabbit
RT skin collagen and its relevance to immunochemical studies of
RT collagen.";
RL -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC (FIBRILLAR FORMING COLLAGEN).
CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
CC -1- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
CC HYDROXYAPATITE.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC PIR: A02856; CGBR15;
DR InterPro: IPR001007; .
DR PROSITE: PS01208; WFCF; PARTIAL.
DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Collagen.
KW MOD_RES 7 7 CONVERTED TO AN ALDEHYDE GROUP THAT IS
FT INVOLVED IN CROSS-LINKING.
FT MOD_RES 26 26 HYDROXYLATION (PROBABLE).
FT MOD_RES 29 29 HYDROXYLATION (PROBABLE).
FT MOD_RES 32 32 HYDROXYLATION (PROBABLE).
FT MOD_RES 41 41 HYDROXYLATION (PROBABLE).
FT MOD_RES 44 44 HYDROXYLATION (PROBABLE).
FT MOD_RES 47 47 HYDROXYLATION (PROBABLE).
FT NON_TER 53 53
SQ SEQUENCE 53 AA; 4987 MW; 127582E5E52B87FC CRC64;

Query Match 11.8%; Score 35; DB 1; Length 53;
Best Local Similarity 30.2%; Pred. No. 1.9e+03;
Matches 13; Conservative 4; Mismatches 24; Indels 2; Gaps 2;
QY 13 VHLPOPVSTRSQHTQPTPSTLPMGPSPAPAGSGTG 55
Db 11 VSVPGMGPSGRGLPGPGAPGPZ-GFZGPPG-2PG2PGSSG 51

RESULT 12
ATP8_SALAL STANDARD; PRT; 55 AA.
AC Q9ANZ1;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
GN MTATP8 OR ATP8 OR ATPASE8.
OS Salvelinus alpinus (Arctic char).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8036;
RN [1]
RP SEQUENCE FROM N.A.
RA Dolron S., Blier P.U., Bernatchez L.;
RA "A comparative analysis of complete sequence of mitochondrial genome
RT between brook char (Salvelinus fontinalis) and arctic char (S.
RT alpinus).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RC FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: AF154851; AAD41389.1; .
DR InterPro: IPR001421; .
DR Pfam: PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ SEQUENCE 55 AA; 6455 MW; 71E430C2E346924A CRC64;
Query Match 11.8%; Score 35; DB 1; Length 55;
Best Local Similarity 30.0%; Pred. No. 2e+03;
Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 11 GAVHLPOPVSTRSQHTQPTP 30
Db 30 GHIFTNEPTSQSTKTRPEP 49
RESULT 13
FOR1_MYRGU STANDARD; PRT; 16 AA.
ID AC P81438;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FORMAEIN 1.
OS Myrmecia gulosa (Red bulldog ant).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Formicoides; Formicidae; Myrmecia.
OX NCBI_TaxID=36170;
RN [1].
RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
RC TISSUE=Hemolymph;
RX MEDLINE=98165787; PubMed=9497332;
RA Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
RT "Isolation from an ant Myrmecia gulosa of two inducible
O-glycosylated proline-rich antibacterial peptides.";
RL J. Biol. Chem. 273:6139-6143(1998).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE. HAS ACTIVITY AGAINST E.COLI
CC BUT NONE AGAINST OTHER GRAM-NEGATIVE BACTERIA AND GRAM-POSITIVE
CC BACTERIA.
CC -1- INDUCTION: UPON BACTERIAL CHALLENGE.
CC -1- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC -1- SIMILARITY: TO DROSOPHILA DROSOCIN.
KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
FT CARBOHYD 11 11 O-LINKED (GALNAC).
FT SEQUENCE 16 AA; 1794 MW; 80CEA3A8B8C2E0AE CRC64;

Query Match 11.6%; Score 34.5; DB 1; Length 16;
Best Local Similarity 41.2%; Pred. No. 6.8e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 5; Gaps 1;
QY 16 PQPVSTRSQHTQPTPEP 32
Db 3 PNPVNK-----PTPHP 14

RESULT 14
ATP8_RHEAM STANDARD; PRT; 55 AA.
ID AC Q79396;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
GN MTATP8 OR ATP8 OR ATPASE8.
OS Rhea americana (Greater rhea) (Common rhea).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Rheiformes; Rheidae; Rhea.
OX NCBI_TaxID=8797;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=98272735; PubMed=9608049;
RA Harlid A., Janke A., Arnason U.;
RT "The complete mitochondrial genome of Rhea americana and early avian
RT divergences";
RL J. Mol. Evol. 46:669-679(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX Sorenson M.D., Dimcheff D.E., Ast J.C., Yuri T., Mindell D.P.;
RA "Complete mitochondrial DNA sequences for five birds and a turtle.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y16884; CAA76505.1; -.
DR EMBL; AF090339; AAD32506.1; -.
DR InterPro; IPR001421; -.
DR InterPro; IPR003237; -.
DR Pfam; PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ SEQUENCE 55 AA; 6493 MW; CF394C4267956648 CRC64;

Query Match 11.4%; Score 34; DB 1; Length 55;
Best Local Similarity 40.7%; Pred. No. 2.4e+03;
Matches 11; Conservative 3; Mismatches 5; Indels 8; Gaps 2;

QY 14 HLP--QPVSTRSQHTQPTPEPSTAPST 38
   ||| :||| ||| |||
DB 35 NLPSNKKPLST-----PNPTWTWPWT 55

RESULT 15
M84C_DROME STANDARD; PRT; 55 AA.
ID AC Q01644; Q9VTA0;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE MALE SPECIFIC SPERM PROTEIN MST84DC.
GN MST84DC OR CG17945.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OREGON-R;
RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
RA Schaefer M.;
RT "A cluster of four genes selectively expressed in the male germ line
RT of Drosophila melanogaster.";
RL Mech. Dev. 35:143-151(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```

```

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.I.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam P.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou G., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- TISSUE SPECIFICITY: TESTIS.
CC -!- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
CC -!- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
CC MOTIFS.
CC -----
CC -!- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67703; CAA47939.1; -.
DR EMBL; AE003672; AAF54025.1; -.
DR HSSP; P01180; INPO.
DR FlyBase; FBgn0004174; Met84DC.
KW Spermatogenesis; Repeat; Multigene family.
SQ SEQUENCE 55 AA; 5225 MW; 95A12F3AEC88BD6C CRC64;

```

```

Query Match 11.4%; Score 34; DB 1; Length 55;
Best Local Similarity 37.5%; Pred. No. 2.4e+03;
Matches 9; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 32 PSTAPSTSFLLPMGSPPPAEGSTG 55
DB 24 PRCGPCGSCCGPCGCGCCGPGF 47

```

Search completed: July 13, 2001, 17:47:36
Job time: 163 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	55.5	18.7	42	4	Q16459	homo sapien	
2	55.5	18.7	52	4	Q14441	homo sapien	
3	52	17.5	52	14	Q92313	human respi	
4	51	17.2	52	14	Q92301	human respi	
5	51	17.2	52	14	Q92314	human respi	
6	50.5	17.0	33	6	Q02832	gorilla gor	
7	50.5	17.0	34	11	Q9WY97	mus musculus	
8	50	16.8	52	6	Q9GL38	bos taurus	
9	49.5	16.7	54	14	Q9J7C8	simian viru	
10	49	16.5	46	4	Q15218	homo sapien	
11	49	16.5	51	6	Q18723	macaca fusc	
12	49	16.5	52	14	Q92302	human respi	
13	49	16.5	52	14	Q92303	human respi	
14	48	16.2	38	4	Q9HB17	homo sapien	
15	48	16.2	42	6	Q9GK73	sus scrofa	
16	47.5	16.0	27	10	Q9SBM0	solanum tub	
17	47	15.8	31	2	Q9Z516	streptomyce	
18	47	15.8	52	14	Q92300	human respi	
19	47	15.8	52	14	Q92311	human respi	

GN GPIB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ishida F.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L39103; AAA69491.1; -
 DR InterPro; IPR002965; -
 DR PRINTS; PRO1217; PRICHEXTENS.
 FT NON_TER 1 1
 FT NON_TER 52 52
 SQ SEQUENCE 52 AA; 5187 MW; 829FBE84792EA30F CRC64;

Query Match 18.7%; Score 55.5; DB 4; Length 52;
 Best Local Similarity 35.6%; Pred. No. 26;
 Matches 16; Conservative 6; Mismatches 18; Indels 5; Gaps 3;

QY 4 PTRSMAGAVHLPQVSTRSHTQTPPEPSTAPSTSFLLPMGPSP 48
 Db 12 PTSEPAFPT-TPEPTSEFAP-SPTTPEPTSEFAPS---PTTPEP 51

RESULT 3
 O92313 PRELIMINARY; PRT; 52 AA.
 AC O92313;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=196775;
 RA Coggins W.B., Lefkowitz E.J., Sullender W.M.;
 RT "Genetic Variability among Group A and Group B Respiratory Syncytial
 RT Viruses in a Children's Hospital."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF086886; AAC43006.1; -
 DR InterPro; IPR000925; -
 DR Pfam; PF00802; Glycoprotein_G; 1.
 FT NON_TER 1 1
 FT NON_TER 52 52
 SQ SEQUENCE 52 AA; 5541 MW; 8C3111FE05DF29E0 CRC64;

Query Match 17.5%; Score 52; DB 14; Length 52;
 Best Local Similarity 34.3%; Pred. No. 62;
 Matches 12; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 14 HLPQVSTRSHTQTPPEPSTAPSTSFLLPMGPSP 48
 Db 13 HTSQKETHLSTSEGNPSPSQVYTTSEYLSQSPSP 47

RESULT 4
 O92301 PRELIMINARY; PRT; 52 AA.
 AC O92301;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=181691;
 RA Coggins W.B., Lefkowitz E.J., Sullender W.M.;
 RT "Genetic Variability among Group A and Group B Respiratory Syncytial
 RT Viruses in a Children's Hospital."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF086872; AAC42992.1; -
 DR InterPro; IPR000925; -
 DR Pfam; PF00802; Glycoprotein_G; 1.
 FT NON_TER 1 1
 FT NON_TER 52 52
 SQ SEQUENCE 52 AA; 5613 MW; 7C3114ACA02574E6 CRC64;

Query Match 17.2%; Score 51; DB 14; Length 52;
 Best Local Similarity 34.3%; Pred. No. 79;
 Matches 12; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 14 HLPQVSTRSHTQTPPEPSTAPSTSFLLPMGPSP 48
 Db 13 HTSQKETHLSTSEGNPSPSQVYTTSEYLSQSPSP 47

RESULT 5
 O92314 PRELIMINARY; PRT; 52 AA.
 AC O92314;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=203721;
 RA Coggins W.B., Lefkowitz E.J., Sullender W.M.;
 RT "Genetic Variability among Group A and Group B Respiratory Syncytial
 RT Viruses in a Children's Hospital."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF086888; AAC43008.1; -
 DR InterPro; IPR000925; -
 DR Pfam; PF00802; Glycoprotein_G; 1.
 FT NON_TER 1 1
 FT NON_TER 52 52
 SQ SEQUENCE 52 AA; 5554 MW; 3B328FBC5B4E9858 CRC64;

Query Match 17.2%; Score 51; DB 14; Length 52;
 Best Local Similarity 31.8%; Pred. No. 79;
 Matches 14; Conservative 5; Mismatches 25; Indels 0; Gaps 0;

QY 5 TRSMAGAVHLPQVSTRSHTQTPPEPSTAPSTSFLLPMGPSP 48
 Db 4 TNSITGNLEHTSQEETHLSTSEGNPSPSQVYTTSEYLSQSPSP 47

RESULT 6
 O02832 PRELIMINARY; PRT; 33 AA.
 ID O02832;
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE HUNTINGTIN PROTEIN (FRAGMENT).
 GN IT15.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.

```

OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326790; PubMed=8766138;
RA Pecheux C., Gall A.L., Kaplan J.C., Dode C.;
RT "Sequence analysis of the CAG triplet repeats region in the Huntington
RL disease gene (IT15) in several mammalian species.";
RL Ann. Genet. 33:81-86(1996).
DR EMBL; S83377; AAB50771.1; -.
FT NON_TER 1
SQ SEQUENCE 33 AA; 3393 MW; 44F558CE5636722A CRC64;

Query Match 17.0%; Score 50.5; DB 6; Length 33;
Best Local Similarity 38.5%; Pred. No. 60;
Matches 15; Conservative 1; Mismatches 12; Indels 11; Gaps 3;

QY 15 LPQPVSTRSQHTQP-TPEPSTAPSTFLLPMGSPSPAEG 52
Db ||||| | | | | | | | | | | | | | | | |
2 LPQP-----PPHGOLLPPQPPQ-----PPPPPPPPPP 30

RESULT 7
Q9WTY7 PRELIMINARY; PRT; 54 AA.
AC Q9WTY7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE ENDOTHELIAL NITRIC OXIDE SYNTHASE (FRAGMENT).
GN NOS3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV129;
RX MEDLINE=90096466; PubMed=9878824;
RA Teichert A.M., Karantzoulis-Fegaras F., Wang Y., Mawji I.A., Bei X.,
RA Gnanapandithen K., Warsden P.A.;
RT "Characterization of the murine endothelial nitric oxide synthase
RT promoter.";
RL Biochim. Biophys. Acta 1443:352-357(1998).
DR EMBL; AF091262; AAD22613.1; -.
FT NON_TER 54
SQ SEQUENCE 54 AA; 5193 MW; 1DA456A21958B2EA CRC64;

Query Match 17.0%; Score 50.5; DB 11; Length 54;
Best Local Similarity 47.8%; Pred. No. 92;
Matches 11; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 28 PTPPESTAPSTFLLPMGSPSPA 50
Db | | | | | | | | | | | | | | | |
34 PAPERQA-----PPSPTRA 51

RESULT 8
Q9GL38 PRELIMINARY; PRT; 52 AA.
AC Q9GL38;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CALPASTATIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Chung H.Y., Davis M.E., Hines H.C.;
RT "PCR-SSCP analysis of the bovine calpastatin gene domain L region.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY008267; AAG23869.1; -.
FT NON_TER 1
FT NON_TER 52
SQ SEQUENCE 52 AA; 5749 MW; CBD5A7449AFDDA89 CRC64;

Query Match 16.8%; Score 50; DB 6; Length 52;
Best Local Similarity 33.3%; Pred. No. 1e+02;
Matches 11; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

QY 4 PTRSMAPCAVHLPPQVSTRSQHTQPTPEPSTAP 36
Db | | | | | | | | | | | | | | | |
19 PKHSSDTGSKHAPKEKAVKSKSEQPPSEKSTKP 51

RESULT 9
Q9J7C8 PRELIMINARY; PRT; 54 AA.
AC Q9J7C8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE LARGE T ANTIGEN (FRAGMENT).
OS Simian virus 40 (SV40).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10633;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MC-028863B-I;
RC Rizzo P., Carbone M.;
RT "SV40 from 1955 commercial parenteral (Salk) poliovaccine.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF180738; AAF28272.1; -.
FT NON_TER 1
SQ SEQUENCE 54 AA; 5868 MW; 556CDA8682C1EFCD CRC64;

Query Match 16.7%; Score 49.5; DB 14; Length 54;
Best Local Similarity 32.7%; Pred. No. 1.2e+02;
Matches 16; Conservative 5; Mismatches 25; Indels 3; Gaps 2;

QY 7 SMAPGAVHLPPQ-VYSTRSQHTQPTPEPSTAPSTFLLPMGSPSPAEGST 54
Db | | | | | | | | | | | | | | | |
8 SQSGSFQAPQPSOSSSXHDHNPYHICRGXTCKKP--PTPPPET 54

RESULT 10
Q15218 PRELIMINARY; PRT; 46 AA.
AC Q15218;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE SALIVARY PROLINE-RICH PROTEIN 2 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298176; PubMed=6089212;
RA Azen E.A., Lyons K.M., McGonigal T., Barrett N.L., Clements L.S.,
RA Maeda N., Vanin E.F., Carlson D.M., Smithies O.;
RT "Clones from the human gene complex coding for salivary proline-rich
RT proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5561-5565(1984).
DR EMBL; K02578; AAA36505.1; -.
FT NON_TER 1
FT NON_TER 46
SQ SEQUENCE 46 AA; 4592 MW; FCE1D38D8DEDC173 CRC64;

```


Query Match 16.5%; Score 49; DB 4; Length 46;
Best Local Similarity 38.2%; Pred. No. 1.2e+02;
Matches 13; Conservative 2; Mismatches 15; Indels 4; Gaps 1;

QY 20 STRSQHTQPTPEPSTAPSTSLFLLPMGPSPPAEGS 53
| | | | | : | | | | | : | | | | | : | | | | | :
Db 1 SARS-----PPRKPPQPPQEGNNGPPPPAGGN 30

RESULT 11
O18723 PRELIMINARY; PRT; 51 AA.
AC O18723
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE COMPLEMENT DECAY-ACCELERATING FACTOR (CD55) (FRAGMENT).
GN DAF.
OS Macaca fuscata fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9543;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RX MEDLINE=98099759; PubMed=9435343;
RA Nonaka M., Nonaka M., Takenaka O., Okada H.;
RT "A new repetitive sequence uniquely present in the decay-accelerating
factor genes";
RL Immunogenetics 47:246-255(1998).
CC -1- FUNCTION: PROTECTION OF CELLS FROM COMPLEMENT-MEDIATED DAMAGE (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
(RCA) FAMILY.
CC EMBL; AB003315; BAA22903.1; -.
KW Complement pathway. 1
FT NON_TER 1 51
FT NON_TER 51 51
SQ SEQUENCE 51 AA; 5545 MW; ED112B05C37548A6 CRC64;

Query Match 16.5%; Score 49; DB 6; Length 51;
Best Local Similarity 28.9%; Pred. No. 1.3e+02;
Matches 13; Conservative 6; Mismatches 22; Indels 4; Gaps 1;

QY 10 PGAVHLPQPVSTRSQHTQPTPEPSTAPSTSLFLLPMGPSPPAEGST 54
| | | | | : | | | | | : | | | | | : | | | | | :
Db 3 PPTVQKPTTVNVRTEVSTPSQKITTPNAQ----ATRTSPASRTT 43

RESULT 12
O92302 PRELIMINARY; PRT; 52 AA.
AC O92302
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=182473;
RA Coggin W.B., Lefkowitz E.J., Sullender W.M.;
RT "Genetic Variability among Group A and Group B Respiratory Syncytial
Viruses in a Children's Hospital";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
EMBL; AF086873; AAC42993.1; -.

DR InterPro; IPR000925; -.
DR Pfam; PF00802; Glycoprotein_G; 1.
FT NON_TER 1
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5530 MW; C2029E4C5B551270 CRC64;

Query Match 16.5%; Score 49; DB 14; Length 52;
Best Local Similarity 31.8%; Pred. No. 1.3e+02;
Matches 14; Conservative 4; Mismatches 26; Indels 0; Gaps 0;

QY 5 TRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTSLFLLPMGPSP 48
| | | | | : | | | | | : | | | | | : | | | | | :
Db 4 TNSTTGNLEHTSQEETHLSTSGNTSPSQAYTTSEYLSQPPSP 47

RESULT 13
O92303 PRELIMINARY; PRT; 52 AA.
AC O92303
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=182701;
RA Coggin W.B., Lefkowitz E.J., Sullender W.M.;
RT "Genetic Variability among Group A and Group B Respiratory Syncytial
Viruses in a Children's Hospital";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF086874; AAC42994.1; -.
DR InterPro; IPR000925; -.
DR Pfam; PF00802; Glycoprotein_G; 1.
FT NON_TER 1 52
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5542 MW; C2028FBC5B551270 CRC64;

Query Match 16.5%; Score 49; DB 14; Length 52;
Best Local Similarity 31.8%; Pred. No. 1.3e+02;
Matches 14; Conservative 4; Mismatches 26; Indels 0; Gaps 0;

QY 5 TRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTSLFLLPMGPSP 48
| | | | | : | | | | | : | | | | | : | | | | | :
Db 4 TNSTTGNLEHTSQEETHLSTSGNTSPSQAYTTSEYLSQPPSP 47

RESULT 14
O9HB17 PRELIMINARY; PRT; 38 AA.
AC O9HB17
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE SIGNALING MOLECULE SPECI BETA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pirone D.M., Fukihsara S., Gutkind S.J., Burbelo P.D.;
RT "SPECS, small binding proteins for CDC42 proteins";
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AF286592; AAG17723.1; -.
SQ SEQUENCE 38 AA; 4245 MW; 3B416F3C5ADF4E91 CRC64;

Query Match 16.2%; Score 48; DB 4; Length 38;
 Best Local Similarity 46.7%; Pred. No. 1.2e+02;
 Matches 14; Conservative 2; Mismatches 6; Indels 8; Gaps 2;

QY 13 VHLPPQVSTRSHTQPTPEPSTAPSTFLL 42
 | | | | | | | | | | | | | | | | | | | | | |
 Db 13 VEKPPQVSL-----PTPHPN--PKSSQLL 34

RESULT 15

Q9GKJ3 PRELIMINARY; PRT; 42 AA.
 AC Q9GKJ3;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE MYOSIN LIGHT CHAIN KINASE (FRAGMENT).
 GN MYLK.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Van Poucke M., Yerle M., Tuggle C., Chardon P., Van Zeveren A.,
 RA Peelman L.J.;
 RT "Integration of porcine chromosome 13 maps."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF222917; AAG41130.1; -.
 KW Kinase.
 FT NON_TER 1 1
 FT NON_TER 42 42
 SQ SEQUENCE 42 AA; 4677 MW; 78BDD867E66EF64F CRC64;

Query Match 16.2%; Score 48; DB 6; Length 42;
 Best Local Similarity 36.6%; Pred. No. 1.4e+02;
 Matches 15; Conservative 2; Mismatches 14; Indels 10; Gaps 2;

QY 13 VHLPPQVSTRS-----QHTQPTPEPSTAPSTFLLPMGPPSP 49
 | | | | | | | | | | | | | | | | | | | | | |
 Db 8 VHSPPQVDFRSVLAKKGTPKTPVPEKLP-----PPKPTTP 42

Search completed: July 13, 2001, 17:47:19
 Job time: 166 sec

6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:41:38 ; Search time 19.93 Seconds

(without alignments)
170.343 Million cell updates/sec

Title: US-09-800-909-2_COPY_202_257

Perfect score: 297

Sequence: 1 STSPTRSMAPGAVHLQPVS.....STSFLLPMGPSPPAEGSTGD 56

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 221992

Minimum DB seq length: 0

Maximum DB seq length: 56

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	23.1	47	21	Human secreted pro
2	66	22.2	49	19	Amino acid sequenc
3	59	19.9	45	12	Linking B region #
4	58	19.5	54	15	Sequence of a pept
5	58	19.5	54	15	Camel Ig 2-heavy c
6	56.5	19.0	41	20	Peptide encoded by
7	55	18.5	46	20	Amphotropic hyperv
8	55	18.5	52	22	Human INTERCEPT 21
9	54.5	18.4	55	21	Human ORFX ORF2454
10	54	18.2	21	9	Sequence of human
11	54	18.2	36	21	Human 5' EST relat

12	53.5	18.0	37	18	AAW26641	H. insoliens family
13	52.5	17.7	21	14	AAW37992	Heavy chain hinge
14	51.5	17.3	50	18	AAW35739	Mucin peptide MUC1
15	51.5	17.3	51	18	AAW31697	Mucin peptide MUC1
16	51	17.2	40	17	AAW88860	Peptide targetted
17	51	17.2	43	20	AAV49096	Amino acid sequenc
18	50	16.8	16	17	AAW99693	IgA protease speci
19	50	16.8	19	20	AAW99681	Human IgA1 hinge r
20	50	16.8	21	20	AAV42923	Core peptide in th
21	50	16.8	25	19	AAW64623	Human IgA1 CH2 reg
22	50	16.8	25	19	AAW61595	Human IgA1 CH2 reg
23	50	16.8	25	20	AAW85774	IgA1 CH2 region pr
24	50	16.8	52	21	AAW85774	Human secreted pro
25	49.5	16.7	40	16	AAW68002	Mucin repeat seque
26	49.5	16.7	40	19	AAW72703	Human mucin peptid
27	49.5	16.7	41	12	AAW15255	Linking B region #
28	49	16.5	27	9	AAW80541	Sequence of portio
29	49	16.5	35	15	AAW49650	Sequence of hinge
30	49	16.5	35	15	AAW49529	Camel 2-chain Ig h
31	49	16.5	46	20	AAW01284	Peptide encoded by
32	49	16.5	47	21	AAW38081	Fragment of human
33	48.5	16.3	34	20	AAW03809	Lck binding active
34	48.5	16.3	40	19	AAW54873	Carcinoma-associat
35	48.5	16.3	53	21	AAW00350	Human secreted pro
36	48	16.2	39	20	AAW74023	Human sucrose isom
37	48	16.2	41	2	AAW10196	Sequence of struct
38	48	16.2	41	20	AAW74094	Gastro-intestinal
39	48	16.2	42	8	AAW71006	Sequence of fragme
40	48	16.2	42	9	AAW80040	Beta-human chorion
41	48	16.2	42	10	AAW91846	Analogue of beta s
42	48	16.2	42	19	AAW69459	HCG antigenic pept
43	48	16.2	42	19	AAW42213	Peptide derived fr
44	48	16.2	42	20	AAW93444	Human hCG beta-sub
45	48	16.2	42	21	AAW20553	Human chorionic go

ALIGNMENTS

RESULT 1

AAW56300

ID AAW56300 standard; Protein; 47 AA.

XX AAW56300;

XX AC AAW56300;

XX DT 13-MAR-2001 (first entry)

XX DE Human secreted protein sequence encoded by gene 99 SEQ ID NO:394.

XX KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
XX KW antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
XX KW cerebroprotective; neutropic; neuroprotective; antibacterial; virucide;
XX KW fungicide; ophthalmological; gene therapy; pathological condition;
XX KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
XX KW neoplasia; cardiovascular disorder; cardiac arrest; cerebral ischemia;
XX KW cerebrovascular disorder; angiogenesis; nervous system disorder;
XX KW Alzheimer's disease; infection; ocular disorder; corneal infection;
XX KW wound healing; skin aging; food additive; preservative.

XX OS Homo sapiens.

XX PN WO200070042-A1.

XX PD 23-NOV-2000.

XX PF 11-MAY-2000; 2000WO-US12788.

XX PR 13-MAY-1999; 99US-0134068.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;
PI Duan RD, Florence KA, Soppet DR;

XX	Camelus dromadarius.
OS	
XX	
PH	Location/Qualifiers
FT	6...40
FT	/note= "hinge region"
FT	1..5
FT	/note= "heavy chain variable region"
FT	41..54
FT	/note= "heavy chain constant region"
XX	
PN	EP584421-A.
XX	
PD	02-MAR-1994.
XX	
PX	21-AUG-1992; 92EP-0402326.
PX	
PR	21-AUG-1992; 92EP-0402326.
XX	(CAST/) CASTERMAN C.
PA	(HAME/) HAMERS R.
PA	
PI	Casterman C, Hamers R;
XX	
DR	WPI; 1994-067061/09.
XX	
PT	New isolated immunoglobulin molecules devoid of light polypeptide chains - consisting of heavy polypeptide chains only, obtd. from Camelid serum, for use as antibodies
PT	
PT	
XX	
PS	Disclosure; Page 19; 35pp; English.
XX	
CC	This 2-chain Ig molecule lacks any light chains. The Ig has e.g. improved solubility and aggregate much less than heavy chains of 4-chain Igs. The Igs can be used normally for e.g. diagnosis,
CC	therapy, in vaccines, for isolation and purification of antigens,
CC	and in the production of anti-idiotypic antibodies.
XX	
SQ	Sequence 54 AA;
	Query Match 19.5%; Score 58; DB 15; Length 54;
	Best Local Similarity 31.6%; Pred. No. 29;
	Matches 12; Conservative 3; Mismatches 23; Indels 0; Gaps
QY	15 LPOPVSTRSQHTOPTPEPSTAPSTFLLPMGSPPAEG 52 :: ::
Db	9 ipdqpkpqdpqpkpqpdkpeptcckcpapellg 46
RESULT 6	
AAY01285	ID AAY01285 standard; Protein; 41 AA.
XX	
AC	AAY01285;
XX	
DT	01-JUN-1999 (first entry)
XX	
DE	Peptide encoded by HRGP gene cassette.
XX	
KW	Synthetic gene; plant; gum; hydroxyproline-rich glycoprotein; HRGP
KW	repetitive proline-rich protein; RPRP; arabinogalactan protein; ARGP
KW	glycopeptide; internal repeat.
XX	
OS	Synthetic.
XX	
PN	WO9903978-A1.
XX	
PD	28-JAN-1999.
XX	
PF	21-JUL-1998; 98WO-US15083.
XX	
PR	20-JUL-1998; 98US-0897556.

XX AAC77446 to AAC77606 encode the proteins given in ABA40237 to ABA43397, CC
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX CC
CC sequences have activities such as: cytostatic; hepatotropic; vulnary; CC
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; CC
CC osteopathic; anticouagulant; antiarthritic; immunosuppressant; CC
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; CC
CC antidiabetic; hypotensive; dermatological; immunosuppressive; CC
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; CC
CC antithyroid; and antianaemic. The sequences can be used for determining

XX AAC77446 to AAC77606 encode the proteins given in ABA40237 to ABA43397, CC
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX CC
CC sequences have activities such as: cytostatic; hepatotropic; vulnary; CC
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; CC
CC osteopathic; anticancerous; antiarthritic; immunosuppressant; CC
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; CC
CC antidiabetic; hypotensive; dermatological; immunosuppressive; CC
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; CC
CC antithyroid; and antianaemic. The sequences can be used for determining

xx
CC A Cu/Zn superoxide dismutase (SOD) polymer of SOD units of at least 2
CC SOD monomers covalently coupled, carboxy terminus to amino terminus, to
CC each other by a polypeptide spacer of at least 3 amino acids is claimed.
CC The polymers are of formula
CC (SOD monomer - Iga^a - SOD monomer)x or
CC (SOD monomer - Iga^a - SOD monomer - Iga^a - SOD monomer)x

CC IgA = a 10-100 amino acid long segment of an immunoglobulin hinge region
 CC x = 2-4
 CC AAN80435 was used to make pSODCF1SODHA1 encoding bacterial expression
 CC spacer-linked human SOD monomers. AAN80435 encodes human IgA1 hinge
 CC region beginning at AA residue 226 to avoid the Cys 225 residue with
 CC BamHI and NcoI sites at its ends. The SOD polymers have an extended
 CC circulatory life and retain the activity of SOD. The cloning and
 CC sequencing of human SOD (hSOD) cDNA and prodn. of hSOD in bacteria and
 CC yeast are described in EP-138111.
 CC
 CC Sequence 21 AA;
 SQ
 Query Match 18.2%; Score 54; DB 9; Length 21;
 Best Local Similarity 48.0%; Pred. No. 28;
 Matches 12; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
 Qy 15 LPQPVSTRSQHTQTPPEPSTAPSTS 39
 Db : ||| : | ||| ||| | :
 1 laqpyps----tpptpsptptps 21
 RESULT 11
 AAY65289
 ID AAY65289 standard; Protein: 36 AA.
 XX AC
 AC AAY65289;
 XX DT 01-FEB-2000 (first entry)
 XX DE Human 5' EST related polypeptide SEQ ID NO:1450.
 XX KW Human: 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification.
 XX OS Homo sapiens.
 XX PN WO9953051-A2.
 XX PD 21-OCT-1999.
 XX PF 09-APR-1999; 99WO-IB00712.
 XX PR 09-APR-1998; 98US-0057719.
 XX PR 28-APR-1998; 98US-0069047.
 XX PA (GEST) GENSET.
 XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-038446/03.
 XX DR N-PSDB; AAZ42903.
 XX PT Novel secreted protein 5' expressed sequence tag sequences used in
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
 XX Claim 3; Page 788; 837pp; English.
 XX AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AAY64651 to
 CC AAY65438 represent the EST-related proteins corresponding to AAZ42265 to
 CC AAZ43052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be

CC used for directing extracellular secretion of a polypeptide or the
 CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AAZ42249 to AAZ42264 and AAY64644 to AAY64650 represent
 CC sequences used in the exemplification of the present invention.
 CC
 CC Sequence 36 AA;
 SQ
 Query Match 18.2%; Score 54; DB 21; Length 36;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 12; Conservative 2; Mismatches 8; Indels 2; Gaps 1;
 Qy 28 PPEPSTAPSTSFLLPMGP--SPP 49
 Db : ||| ||| : | : |||
 9 pspapsalappsfalslppsapp 32
 RESULT 12
 AAW26641
 ID AAW26641 standard; Peptide: 37 AA.
 XX AC
 AC AAW26641;
 XX DT 08-FEB-1998 (first entry)
 XX DE H. insolens family 45 cellulase linker peptide.
 XX KW Desizing; cellulose; fabric; enzyme hybrid; cellulase;
 KW Myceliophthora thermophila; cellulose binding domain; lipase;
 KW Lipolase; Humicola insolens; linker.
 XX OS Synthetic.
 XX PN WO9728256-A1.
 XX PD 07-AUG-1997.
 XX PF 29-JAN-1997; 97WO-DK00041.
 XX PR 29-JAN-1996; 96DK-0000093.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX PI Bjornvad ME, Rasmussen MD, Vind J, Von Der Osten C;
 XX WPI; 1997-402610/37.
 XX PT Desizing cellulose-containing fabric or textile using an enzyme
 PT hybrid - which comprises a catalytically active amino acid sequence
 PT of a non-cellulolytic enzyme linked to an amino acid sequence
 PT comprising a cellulose binding domain.
 XX Example 4; Page 43; 72pp; English.
 XX This peptide comprises a Humicola insolens family 45 cellulase
 CC glycosylated linker used in a novel hybrid enzyme between the
 CC cellulose binding domain (CBD) of Myceliophthora thermophila
 CC cellulase and Humicola lanuginosa lipase (Lipolase). DNA
 CC expressing the linker was obtained by PCR (see AAT90506-07). A DNA
 CC construct encoding CBD-glycosylated linker-lipolase is given in
 CC AAT7070. A claimed process for desizing cellulose-containing fabric
 CC or textile involves treating the fabric or textile with a modified
 CC enzyme (enzyme hybrid) comprising a non-cellulolytic enzyme linked
 CC to a CBD. The process gives improved enzyme performance by
 CC modifying the enzyme so as to increase its affinity for cellulosic
 CC fabric.
 XX Sequence 37 AA;
 SQ


```
Query Match      18.0%; Score 53.5; DB 18; Length 37;
Best Local Similarity 36.8%; Pred. No. 55;
Matches 14; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 13 VHLPPQVSTRSQHTQPTPEPSTAPSTSFLLPMGPSPPA 50
   | | | | | | | | | | | | | | | | | | | |
Db 1 vqipss-stsgpvnqptstststststssppvqpttps 37

RESULT 13
ID AAR37992 standard; peptide; 21 AA.
XX AC AAR37992;
XX DT 13-OCT-1993 (first entry)
XX DE Heavy chain hinge region of IgA1.
XX KW Heavy chain; hinge region; human; IgA1; IgA; secretion; adult; IgA2;
XX KW newborn; children; milk; infant; passive; immunity; mucous membrane;
XX KW antibody; bacteria; IgA protease; allotype; resistant.
XX OS Homo sapiens.
XX PN W09310818-A.
XX PD 10-JUN-1993.
XX PF 03-DEC-1992; 92WO-US10432.
XX PR 04-DEC-1991; 91US-0802338.
XX PA (NEW-) NEW ENGLAND MED CENT HOSPITALS.
XX PI Plaut AG;
XX DR WPT; 1993-196741/24.
XX PT Infant formula contg. an IgA protease inhibitor - for providing
XX PT protection against IgA protease-producing pathogens
XX PS Disclosure; Page 17; 25pp; English.
XX CC This sequence represents the heavy chain hinge region of human IgA1.
XX CC IgA is secreted by adults at high concentrations in their secretions.
XX CC However newborn children have no mucosal IgA of their own. Human
XX CC milk is very rich in IgA providing the infant with passive immunity
XX CC by coating the mucous membranes with the antibody. Some bacteria
XX CC secrete IgA proteases which inactivate the infants IgA, when it starts
XX CC producing it, by cleavage at the hinge region, leading to infection of
XX CC the infant. The most common allotype of IgA2 has the hinge region
XX CC deleted and is therefore resistant to degradation by bacterial
XX CC proteases. See also AAR37993.
XX SQ Sequence 21 AA;

Query Match      17.7%; Score 52.5; DB 14; Length 21;
Best Local Similarity 45.2%; Pred. No. 39;
Matches 14; Conservative 1; Mismatches 3; Indels 13; Gaps 2;

QY 18 PVSTRSQHTQPTPEPSTAPSTSFLLPMGPSP 48
   | | | | | | | | | | | | | |
Db 2 pvs-----tpttpspstpt-----psp 19

RESULT 14
AAW35739
ID AAW35739 standard; peptide; 50 AA.
XX AC AAW35739;

Query Match      17.3%; Score 51.5; DB 18; Length 50;
Best Local Similarity 31.5%; Pred. No. 1.2e+02;
Matches 17; Conservative 6; Mismatches 18; Indels 13; Gaps 2;

QY 1 STSPTRSMAPGAVHLPPQVSTRSQHTQPTPEPSTAPSTSFLLPMGPSPPAEGST 54
   : : : | | | | | | | | | | | | | | | |
Db 2 tsapdtrpapgstappahgvtspaptrpap-gsta-----ppahgvt 42

RESULT 15
AAW31697
ID AAW31697 standard; peptide; 51 AA.
XX AC AAW31697;
XX DT 02-APR-1998 (first entry)
XX DE Mucin peptide MUC1-5 with N-terminal cysteine.
XX KW MUC1; mucin; vaccine; keyhole limpet haemocyanin; KHL; cancer;
XX KW immunogenic protein; immune response; conjugate.
XX OS Synthetic.
XX OS Mammalia.
XX PN W09734921-A1.
XX PD 25-SEP-1997.
```

```
XX 20-MAR-1997; 97WO-US04493.
XX
XX 20-MAR-1996; 96US-0013775.
XX
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
XX Livingston PO, Zhang S;
XX
XX WPI; 1997-480158/44.
XX
XX Vaccine effective against cancer of the breast, prostate, colon,
XX lung or pancreas - comprising mucin peptide, especially MUC1,
XX conjugated to immunogenic protein, especially keyhole limpet
XX haemocyanin
XX
XX Disclosure; Page 11; 45pp; English.
XX
XX This mucin peptide is used in a vaccine capable of producing an immune
XX response which recognises a mucin. The mucin peptide is selected from
XX MUC1 peptide group. The vaccine comprises an amount of the mucin peptide
XX conjugated to an immunogenic protein effective to stimulate or enhance
XX immune response in the subject, together with an adjuvant and a vehicle.
XX A cysteine is added to the N-terminal of this peptide to facilitate the
XX conjugation with protein carriers. The immunogenic protein is a keyhole
XX limpet haemocyanin (KLH) or its derivative. The vaccine can be used to
XX induce an immune response in patients suffering from a cancer of the
XX type where the cancer cells have mucin on their surface, e.g. breast
XX cancer, prostate cancer, lung cancer, colon cancer or pancreas cancer.
XX
SQ Sequence 51 AA;
```

Query Match 17.3%; Score 51.5; DB 18; Length 51;
Best Local Similarity 31.5%; Pred. No. 1.2e+02;
Matches 17; Conservative 6; Mismatches 18; Indels 13; Gaps 2;

```
QY 1 STSPTRSMAPGAVHLPQVSPRSQHTOPTPEPSTAPSTSFLLPMGCPSPPEGST 54
   ::|||::| | | | | | | | | | | | | | | | | | | | | |
Db 3 tsapdtrpapgstappahgvtspap-gsta-----ppahgvt 43
```

Search completed: July 13, 2001, 17:44:30
Job time: 172 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2001, 17:42:28 ; Search time 12.1 Seconds
(without alignments)
93.231 Million cell updates/sec

Title: US-09-800-909-2_COPY_202_257

Perfect score: 297

Sequence: 1 STSPTSMAPCAVHLQPVS.....STSFLLPMGSPPAEGSTGD 56

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 2014635 residues

Total number of hits satisfying chosen parameters: 139014

Minimum DB seq length: 0

Maximum DB seq length: 56

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	27.9	15	US-08-221-583-56	Sequence 56, Appl
2	83	27.9	15	PCT-US95-04018-56	Sequence 56, Appl
3	82	27.6	15	US-08-221-583-58	Sequence 58, Appl
4	82	27.6	15	PCT-US95-04018-58	Sequence 58, Appl
5	81	27.3	15	US-08-221-583-57	Sequence 57, Appl
6	81	27.3	15	US-08-221-583-62	Sequence 62, Appl
7	81	27.3	15	PCT-US95-04018-57	Sequence 57, Appl
8	81	27.3	15	PCT-US95-04018-62	Sequence 62, Appl
9	80	26.9	15	US-08-221-583-55	Sequence 55, Appl
10	80	26.9	15	US-08-221-583-59	Sequence 59, Appl
11	80	26.9	15	PCT-US95-04018-55	Sequence 55, Appl
12	80	26.9	15	PCT-US95-04018-59	Sequence 59, Appl
13	79	26.6	15	US-08-221-583-61	Sequence 61, Appl
14	79	26.6	15	PCT-US95-04018-61	Sequence 61, Appl
15	78	26.3	15	US-08-221-583-60	Sequence 60, Appl
16	78	26.3	15	PCT-US95-04018-60	Sequence 60, Appl
17	64	21.5	15	US-08-221-583-54	Sequence 54, Appl
18	64	21.5	15	PCT-US95-04018-54	Sequence 54, Appl
19	59	19.9	45	US-08-361-920-19	Sequence 19, Appl
20	59	19.9	45	US-08-479-939-19	Sequence 19, Appl
21	59	19.9	45	US-08-483-432-19	Sequence 19, Appl
22	58	19.5	54	US-08-471-780C-44	Sequence 44, Appl
23	58	19.5	54	US-08-467-282B-44	Sequence 44, Appl
24	58	19.5	54	US-08-471-282A-44	Sequence 44, Appl
25	58	19.5	54	US-08-466-710C-44	Sequence 44, Appl
26	58	19.5	54	US-08-468-739C-44	Sequence 44, Appl
27	55	18.5	46	US-08-856-074A-39	Sequence 39, Appl

28	55	18.5	48	6	5171685-7	Patent No. 5171685
29	55	18.5	48	6	5518916-7	Patent No. 5518916
30	53.5	18.0	37	3	US-08-814-052-37	Sequence 37, Appl
31	53.5	18.0	37	3	US-08-812-829-29	Sequence 29, Appl
32	52.5	17.7	21	5	PCT-US92-10432-1	Sequence 1, Appl
33	52.5	17.7	22	1	US-08-442-542-43	Sequence 43, Appl
34	52.5	17.7	22	3	US-08-765-469-43	Sequence 43, Appl
35	50.5	17.0	53	6	5422248-4	Patent No. 5422248
36	50	16.8	15	1	US-08-221-583-53	Sequence 53, Appl
37	50	16.8	15	5	PCT-US95-04018-53	Sequence 53, Appl
38	50	16.8	16	1	US-08-366-591-15	Sequence 15, Appl
39	49.5	16.7	33	1	US-08-237-716-11	Sequence 11, Appl
40	49.5	16.7	40	1	US-08-099-354-1	Sequence 1, Appl
41	49.5	16.7	40	2	US-08-288-059-7	Sequence 7, Appl
42	49.5	16.7	41	1	US-08-361-920-13	Sequence 13, Appl
43	49.5	16.7	41	1	US-08-479-939-13	Sequence 13, Appl
44	49.5	16.7	41	1	US-08-483-432-13	Sequence 13, Appl
45	49	16.5	16	1	US-08-471-033-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-08-221-583-56
; Sequence 56, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-221-583-56

Query Match 27.9%; Score 83; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 HLPQPVSTRSQHTQP 28
|||||
Db 1 HLPQPVSTRSQHTQP 15

RESULT 2

PCT-US95-04018-56
; Sequence 56, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-56

Query Match 27.9%; Score 83; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 HLPQPVSTRSQHTOP 28
Db 1 HLPQPVSTRSQHTOP 15
|||||

RESULT 3
US-08-221-583-58
; Sequence 58, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania

COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25.mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-58

Query Match 27.6%; Score 82; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 SQHTQTPTEPTAPS 37
Db 1 SQHTQTPTEPTAPS 15
|||||

RESULT 4
PCT-US95-04018-58
; Sequence 58, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:

NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0232
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-04018-58

Query Match 27.3%; Score 82; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 SQTQTPPEPSTAPS 37
|||||
DB 1 SQTQTPPEPSTAPS 15

RESULT 5
US-08-221-583-57
Sequence 57, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
CURRENT APPLICATION DATA:
FILING DATE:
APPLICATION NUMBER: US/08/221,583
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-57

Query Match 27.3%; Score 81; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 STRSQTQTPPEPST 34
|||||
DB 1 STRSQTQTPPEPST 15

RESULT 6
US-08-221-583-62
Sequence 62, Application US/08221583
Patent No. 5486595
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CCOR-0185
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-62

Query Match 27.3%; Score 81; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 APSTSFLLPMGPSP 49
|||||
DB 1 APSTSFLLPMGPSP 15

RESULT 7
PCT-US95-04018-57
Sequence 57, Application PC/TUS9504018
GENERAL INFORMATION:
APPLICANT: Heavner, George A.
APPLICANT: Kruszynski, Marian
APPLICANT: Mervic, Miljenko
APPLICANT: Weber, Robert W.
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: Norris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-04018-57

Query Match 27.3%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 STRSQHTQTPPEST 34
| | | | | | | | | | | | | | | |
Db 1 STRSQHTQTPPEST 15

RESULT 8
PCT-US95-04018-62
; Sequence 62, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESS: Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994

; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-04018-62

Query Match 27.3%; Score 81; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 APSTSFLLPMGSPSP 49
| | | | | | | | | | | | | | | |
Db 1 APSTSFLLPMGSPSP 15

RESULT 9
US-08-221-583-55
; Sequence 55, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-221-583-55

Query Match 26.9%; Score 80; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 GAVHLPQPVSTRSQH 25
| | | | | | | | | | | | | | | |
Db 1 GAVHLPQPVSTRSQH 15

RESULT 10
US-08-221-583-59
; Sequence 59, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25:mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221.583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-59

Query Match 26.9%; Score 80; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 TQTPPEPSTPSTSF 40
Db 1 TQTPPEPSTPSTSF 15
|||||

RESULT 11
PCT-US95-04018-55
; Sequence 55, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-55

Query Match 26.9%; Score 80; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 GAVHLPQPVSTRSQH 25
Db 1 GAVHLPQPVSTRSQH 15
|||||

RESULT 12
PCT-US95-04018-59
; Sequence 59, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581

;; FILING DATE: 01-APR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Deluca, Mark
;; REGISTRATION NUMBER: 33,229
;; REFERENCE/DOCKET NUMBER: CCOR-0232
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-3100
;; TELEFAX: (215) 568-3439
;; INFORMATION FOR SEQ ID NO: 59:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US95-04018-59

Query Match 26.9%; Score 80; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 TQTPPEPSTAPSTSF 40
Db 1 TQTPPEPSTAPSTSF 15

RESULT 13
US-08-221-583-61
; Sequence 61, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25.mdtchMod.
;; CURRENT APPLICATION DATA:
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Deluca, Mark
;; REGISTRATION NUMBER: 33,229
;; REFERENCE/DOCKET NUMBER: CCOR-0185
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-3100
;; TELEFAX: (215) 568-3439
;; INFORMATION FOR SEQ ID NO: 61:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-221-583-61

Query Match 26.6%; Score 79; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 PSTAPSTSFLLPMGP 46
Db 1 PSTAPSTSFLLPMGP 15

RESULT 14
PCT-US95-04018-61
; Sequence 61, Application PC/TUS9504018
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; APPLICANT: Mervic, Miljenko
; APPLICANT: Weber, Robert W.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04018
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,580
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,583
; FILING DATE: 01-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,581
; FILING DATE: 01-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04018-61

Query Match 26.6%; Score 79; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 PSTAPSTSFLLPMGP 46
Db 1 PSTAPSTSFLLPMGP 15

RESULT 15
US-08-221-583-60
; Sequence 60, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris

STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdtcMod.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,583
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CGOR-0185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-221-583-60

Query Match 26.3%; Score 78; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 TPEPSTAPSTSELLP 43
|||||
Db 1 TPEPSTAPSTSELLP 15

Search completed: July 13, 2001, 17:44:49
Job time: 141 sec